INTRODUCTION TO RESEARCH IN EDUCATION

مقدمة للبحث في التربية

تأليف

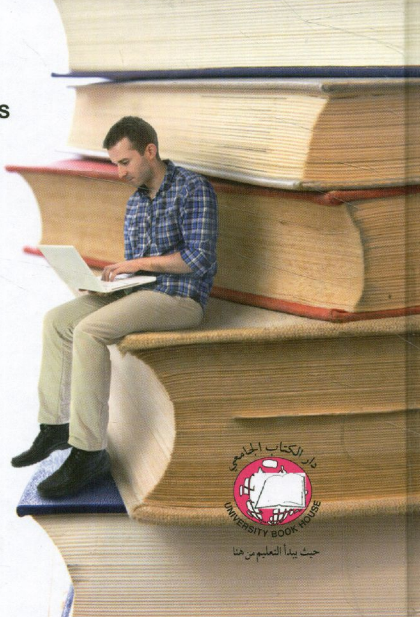
DONALD ARY

LUCY CHESER JACOBS

ASGHAR RAZAVIEH

ترجمة

سعـــد الحسينـــي







شركة جمال أحمد محمد حيف وإخوانه
www.massira.jo



شركة جمال أحمد محمد حيف وإخوانه
www.massira.jo



مقدمة للبحث في التربية

INTRODUCTION TO RESEARCH IN EDUCATION

الطبعــة الأولــى 2013م – 1434هـ

دار الكتاب الجامعي
العين . الإمارات العربية المتحدة
ص.ب . ١٦٩٨٣ - فاكس ٧٥٤٢١٠٢
هاتف ٧٥٥٤٨٤٥ - ٧٥٥٦٩١١ (٣) (٩٧١)

UNIVERSITY BOOK HOUSE
Al Ain - United Arab Emirates
P.O.Box 16983- Fax:75 42102
Tel:(971) (3)7554845- 7556911
bookhous@emirates.net.ae -WWW.bookhous.com-tboourji@yahoo.com



شركة جمال أحمد محمد حيف وإخوانه

عنوان الدار

الرئيسي : عمان - العبدلي - مقابل البنك العربي هاتف : 5627049 6 962+ فاكس : 5627059 6 962+
الفرع : عمان - ساحة المسجد الحسيني - سوق البتراء هاتف : 4640950 6 962+ فاكس : 4617640 6 962+
صندوق بريد 7218 عمان - 11118 الأردن

E-mail: info@massira.jo . Website: www.massira.jo

مقدمة للبحث في التربية

INTRODUCTION TO RESEARCH IN EDUCATION

تأليف

DONALD ARY

LUCY CHESER JACOBS

ASGHAR RAZAVIEH

ترجمة

سعد الحسيني

مراجعة

الدكتور عادل عبدالكريم ياسين



هذه ترجمة عربية مصرح بها لكتاب

INTRODUCTION TO RESEARCH IN EDUCATION

Fifth Edition

للمؤلفين

Donald Ary
(Northern Illinois University)

Lucy Cheser Jacobs
(Indiana University)

Asghar Razavieh
(Shiraz University, Shiraz, Iran)

Harcourt Brace College Publishers



الموضـــوع	الصفحة

الجزء الثاني خلفية البحث

الفصل الثاني: مشكلة البحث

الفصل الثالث: مراجعة الأدبيات

الفصل الرابع: الفرضية

الجزء الثالث: التحليل الإحصائي

الفصل الخامس: الإحصاء الوصفي

الفصل السادس: المعاينة والاستدلال الإحصائي

الجزء الرابع: أسس القياس

الفصل السابع: أدوات البحث

الفصل الثامن: الصدق / الصحة والثبات / الاستقرار

الجزء الخامس: طرائق البحث

الفصل التاسع : البحث التجريبي

الفصل العاشر: البحث العلّي - المقارن

الفصل الحادي عشر: البحث الارتباطي

الفصل الثاني عشر - البحث المسحي

الفصل الثالث عشر - البحث النوعي والتاريخي

الجزء السادس: إيصال البحث

الفصل الرابع عشر: الخطوط العريضة لكتابة مقترحات البحوث

الفصل الخامس عشر: تحليل، وتفسير، وكتابة النتائج



تصدير

"اليقين في المعرفة الوضعية – التي يصنعها البشر – تظل رحلة إبحار لا منتهية في بحر الشك، تقلع من ميناء للشك لترسو في ميناء آخر للشك".

لقد جذبتني الطبعة الأولى لهذا الكتاب "مقدمة للبحث في التربية" إبّان دراستي للدكتوراه، مما جعلني أقرره كمرجع رئيس خلال التسعينات من القرن المنصرم لطلبتي في كلية الدراسات العليا بجامعة أوتارا ماليزيا (UUM) في ماليزيا. وقد لمست آثاره المتميزة في تقدم فهمهم للبحث العلمي في التربية، وتمنيت، حقاً، أن أرى له ترجمة في اللغة العربية. ولربما كان لهذا الإعجاب أثره بقبول هذا العمل الشاق لمراجعة ترجمته.

إن هذا الكتاب مرجع متميز، بالرغم من أنه يحمل كلمة "مقدمة" في عنوانه، وهذا دأب كثير من العلماء في التواضع المحمود، ومثالنا كتاب "مقدمة" ابن خلدون الذي خطّ، بكتابه هذا، النهج لفهم التاريخ والعلوم الاجتماعية.

يتميز هذا الكتاب بالثراء في الأمثلة التي يطرحها لفهم الأفكار التي يتحدث عنها، وهذا يتيح للباحث المبتدئ إدراك حقيقة وحدود الأفكار التي يتصدى لمعالجتها. ويحتوي كل فصل في الكتاب على مجموعة من الأسئلة المنتقاة لتعميق التبصّر في المفاهيم والطرق والمبادئ التي يعالجها الفصل، كما أن هناك إجابات واضحة لكل سؤال يعرضه بما يسهّل على المعلم والطالب التمكّن من هذه الأفكار. ويبرز الكتاب قائمة بالمفاهيم الأساسية الواردة في كل فصل.

لقد تمسّكنا بعرض الاسم الأجنبي للعلماء ودراساتهم ليتسنى للمهتمين العودة إليها إن شاءوا، إذ أن عرض الاسم باللغة العربية قد لا يتيح للباحث المهتم العودة إلى دراسات بحوث العلماء المعنيين فذلك أمر صعب المنال.

وتلافيا لوقوع أخطاء في تعريب الجداول والمعادلات، فقد احتفظنا بالجداول والمعادلات حسب صيغتها الواردة في الكتاب توخيا لدقة المراجعة. وييسّر هذا عملية الربط بين الرمز المقترح مع مثيله في الكتب الأجنبية (الإنجليزية) التي تلتزم بوحدة الرمز والمصطلح.

وفي ترجمتي للمسرد – قائمة شرح المصطلحات – Glossary، حافظت على توحيد ترجمة هذا المصطلح كما هو وراد في المتن. فقد ترجمت "Norm-reperencedtest" إلى "اختبار مرجعي المتوسط (الوسط/الوسيط)" ذلك لأن المعيار هنا هو "المتوسط" وكذلك ترجمت

"Criterion-referenced test" إلى "اختبار مرجعي الإتقان" لأن "المحك" – وهو معيار أيضاً – يشير إلى "الإتقان" في هذا الاختبار. فالترجمة، كما القمر، تظل "مجرد حجر كبير فحسب" إن فقدت "هالته" التي تشير هنا إلى الدلالة والوضوح والترابط. فالكلمات توجد لتوضيح المعاني والدلالات لا لكي تحجبها.

وبهذه المناسبة، أتمنى أن نرى مؤسسة عربية، تعني بالصالح العام، تتوجه لدعم فكرة الدعوة لتبني مجموعة من المهتمين في نقل المعرفة إلى العربية لإعادة النظر في ترجمة المصطلحات، لتطوير ما يلزم منها، خدمة للناشئة والمعرفة، وقد يكون ذلك عملاً بارزاً في توحيد المصطلحات في عالمنا العربي.

لقد قمت بمراجعة ترجمة الكتاب، كلمة – كلمة وجملة – جملة، لتأكيد ما اجتهدت أنه نقل أمين وواضح للنص الأصلي. ومن ثم، فإن الأمانة تقتضي أن أكون الملوم الوحيد لأي تقصير أو عيوب في صورة هذا النص النهائي للترجمة، وأي استحسان للترجمة يظل من نصيب المترجم ونصيبي!!.

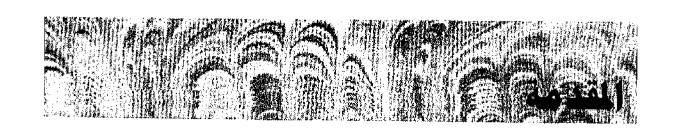
وأشكر بهذه المناسبة جهود **"دار الكتاب الجامعي"** في نشر الكتب المترجمة طالما تتوفر لها ترجمة دقيقة أمينة واضحة، إذ أظن، وقد يكون بعض الظن إثماً، أننا لا نزال بحاجة واسعة للترجمة والنقل في ميادين كثيرة، قد لا يكون "التأليف في مضمارها" قد نضج بعد.

وختاماً أتوجه بالشكر **للدكتور أحمد صالح الألوسي** من كلية التربية بجامعة عجمان لإعانتي في تثبيت ترجمة بعض المصطلحات الإحصائية لباعه الطويل في هذا المجال، كما أشكر **الدكتور مازن أحمد جرادات** من كلية التربية بجامعة عجمان لإعانتي في تثبيت وإصلاح ما أُرتج عليّ في ضبط قواعد كتابة بعض الكلمات العربية، ألم يقل العلاّمة الأستاذ سعيد الأفغاني أنه قد يموت وفي نفسه "غصّة" من تثبيت قواعد "الهمزة"، يرحمه الله.

وأشكر أخيراً **الأستاذ عادل محمود نـــدا** الذي واجه حتما صعوبات تصحيحاتي لأربع مراجعات تصحيح، وتدقيق، لهذه الترجمة، فلولا مجهوده المضني والمتميز في تنسيق وإخراج الكتاب ومثابرته وسعة صدره لما وجد هذا الكتاب سبيله على هذا النحو، الذي أدعو له من الله أن يكون مُرْضياً للقارئ.

المراجـع

د. عادل عبدالكريم ياسين
جامعة عجمان - كلية التربية
2003/10/21



لقد كان هدفنا أثناء تهيئة الطبعة الأولى من (مقدمة للبحث في التربية) في سنة 1972 هو الخروج بكتاب يتمكن القراء بواسطته إتقان الكفايات الضرورية إلى (أولاً) فهم تقييم بحوث آخرين، و (ثانياً) التخطيط ومن ثم القيام ببحثٍ أصيل بمعونة ضئيلة. إن القبول المستمر الذي حَظِي به هذا الكتاب من خلال أربع طبعات سابقة يُشير إلى النجاح في الوصول إلى ذلك الهدف.

وفي الوقت الذي كان فيه البحث في التربية قد تطوّر ونضج، فقد قمنا بتعديل كل طبعةٍ جديدة من هذا الكتاب حتى نقدم إلى قرائنا المعرفة اللازمة التي يحتاجونها لمواكبة هذه التغيرات.

إن تسلسل الموضوعات التي جرى بحثها ومناقشتها في هذا الكتاب يبدأ بوصفٍ عام للطريقة العلمية ووثاقة صلة هذه الطريقة في البحث عن المعرفة. ونحن نفترض بأن قراءنا ليسوا على اطلاع بمفاهيم، وافتراضات، ومصطلحات الطريقة العلمية، ومن ثم جرى توضيح معانيها عند عرضها. ويعقب ذلك اقتراحات بترجمة مشكلات عامة إلى أسئلة محددة قابلة للاستقصاء العلمي من خلال تحديد المجتمع الإحصائي والمتغيرات المعنية.

وبعد ذلك، نقوم بوصف دور البحوث السابقة والجانب النظري في تخطيط مشروع بحث. وقد حدّثنا مصادر الأدبيات ذات الصلة، وأكّدنا بوجه خاص على قواعد البيانات التي تتيح لنا سبيلاً فاعلاً للبحوث ذات الصلة في الجانب النظري وكيفية الوصول إليها.

وتابعنا بعدئذ استقصاء الطرق التي من خلالها تقود كل من النظرية، والخبرة، والملاحظات، والأدبيات ذات الصلة إلى تكوين الفرضيات.

ولقد جرى وصف أساليب المعاينة وأساليب الإحصاء الوصفي والاستدلالي الأوسع استخداما مع التأكيد على دورها في طريقة البحث وعلى تفسيراتها. وجرى استطلاع دور الملاحظة المنتظمة والقياس، كما تضمن أمثلة لأساليب قياس مفيدة. وقد أعيد تنقيح الفصل الخاص بالثبات والصدق، وتحديثه وتوسيعه.

وبعد ذلك، ناقشنا الأنماط المختلفة للبحوث والتي برهنت على فائدتها في التربية، كما أشرنا إلى مزايا ومساوئ مختلف الطرق دون التحيز إلى واحدة بالذات على أساس كونها الأفضل بين الأخريات. وفي هذه الطبعة، قمنا بتوسعة الجزء الخاص بالبحث النوعي، كما قمنا بتوسعة الجزء الخاص بالبحث الارتباطي في فصل منفصل.

واختتمنا الطبعة بتقديم القواعد العامة لتفسير نتائج البحث والأساليب المقبولة لعرض تلك النتائج. وقمنا بتحديث الجزء الخاص بالاعتبارات القانونية، كما قمنا بإضافة مسرد يشرح المصطلحات الواردة.

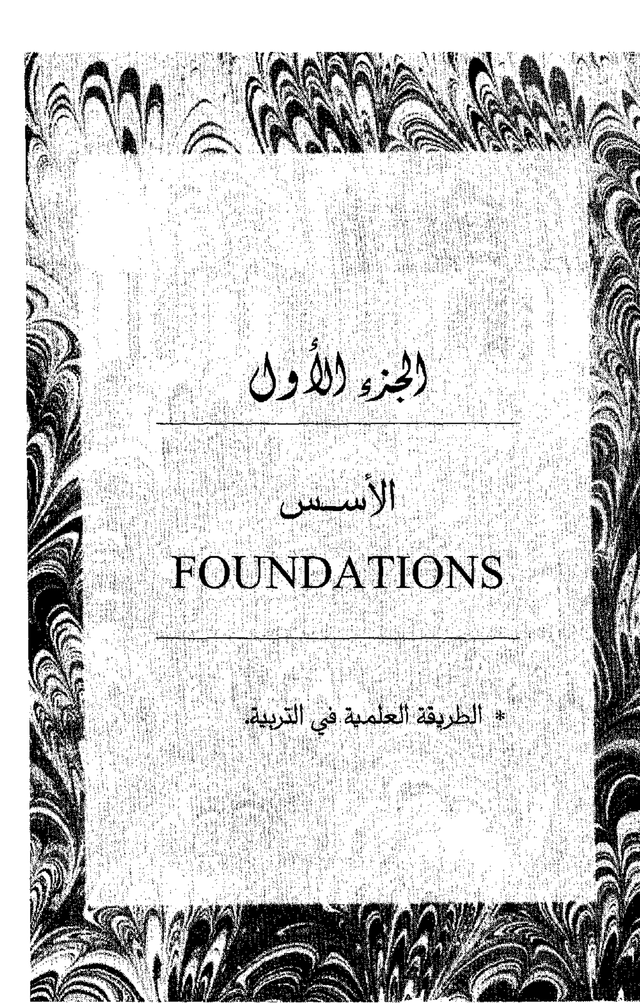
ويبقى مركز الاهتمام بهذه الطبعة هو تقديم كتاب مصمم للاستخدام في مقرر دراسي تمهيدي في مجال البحث التربوي. والهدف منه هو تهيئة إطار لتقييم البحوث المنشورة وكذلك تعريف الباحث المبتدئ بالإجراءات المتبعة للقيام بمشروع بحثي أصيل. ونقوم بالتركيز على المشكلات العلمية النموذجية التي تواجهها البحوث مبتدئين بصياغة أسئلة ومستمرين حتى الإعداد للتقرير النهائي.

وبالإضافة إلى الأسئلة الدراسية الموجودة في نهاية كل فصل من الفصول، ثمة أسئلة مشتملة على أسئلة متعددة الخيارات يجدها القارئ في الكتيب الوجيز الخاص بالمدرس.

وبالرغم من أن الكتاب (مقدمة للبحث في التربية) موجه إلى الطالب المبتدئ في البحث التربوي، فإنا نأمل أن يجده، من تتوفر لديهم رغبة في الاستزادة من الإطلاع على فلسفة، وأدوات، وطرائق البحث العلمي، مفيداً لهم.

وكان المعيار الأساسي المستخدم فيما يجب أن يتضمنه الكتاب، الفائدة المحتملة للأوجه المتنوعة للبحث التربوي للشخص الممارس التربوي.

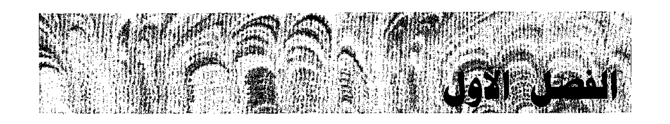
وإلى جميع أولئك المدرسين، الذين استخدموا الطبعات الأربع الأولى، الذين قدموا اقتراحات قيّمة جداً لتحسين ولتحديث الطبعة الخامسة، نقدم لهم شكرنا العميق لجميلهم. كما نشكر السيد (صموئيل هوانك) لعمله الدؤوب الشامل في تقديمه فصلاً شاملاً وحديثاً في مراجعة الأدبيات ذات الصلة. كما نشكر أيضا السيد (وزلي كوفالت) لعمله في الكتيب الوجيز الخاص بالمدرس.



الجزء الأول

الأسس
FOUNDATIONS

* الطريقة العلمية في التربية.



الطريقة العلمية في التربية
The Scientific Approach in Education

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون بوسع الطالب:

1- كتابة المصادر الخمسة الرئيسة للمعرفة، كما سيقوم بالتعليق على كل منها.
2- وصف مميزات الطريقة العلمية.
3- ذكر الافتراضات التي تشكل الأساس للعلم، وكذلك الاتجاهات المميزة للعلماء.
4- تحديد الغاية والمميزات للنظرية العلمية في العلوم السلوكية.
5- الإشارة إلى القصور المتضمن في تطبيق الطريقة العلمية في العلوم الاجتماعية.
6- التمييز بين مميزات البحث الأساسي والبحث التطبيقي.
7- توضيح وتفسير المصطلحات: المفهوم، المفهوم البنائي، والمتغير.
8- التمييز بين الأنماط المتنوعة للمتغيرات: النوعي مقابل المستمر، المستقل مقابل التابع، والفعال / النشيط مقابل المنسوب.
9- وصف نماذج متنوعة من التعاريف: تأسيسي مقابل إجرائي، وتجريبي مقاس مقابل تجريبي إجرائي.
10- وصف طرائق البحث الكبرى المستعملة في البحوث التربوية.

التربويون، بسبب الحاجة، صانعو قرارات. فنحن يوميا، أثناء القيام بأعباء العملية التربوية نواجه مهمة اتخاذ القرارات حول كيفية التخطيط لخبرات التعلم، وكيفية التدريس، وكيفية إرشاد الطلبة، وكيفية إدارة أي نظام مدرسي، وأمور أخرى لا تعد ولا تحصى.

ويجب أن يقوم المهنيون بالتخطيط لأنفسهم، خلافاً للعمال غير الماهرين الذين يقال لهم ما يجب أن يقوموا به من واجبات وكيفية القيام بها. ويفترض أن أولئك المهنيين يمتلكون المعرفة والمهارات الضرورية لاتخاذ قرارات صائبة. لكن أنى للتربويين أن يعلموا ما هو الجواب الصحيح في موقف محدد؟ وبالرغم من وجود مصادر أخرى للمعرفة مثل الخبرة، والمرجعية، والتقاليد، إلا أن المعرفة العلمية حول العملية التربوية هي التي تقدم العطاء الأهم في صناعة القرار في مجال التربية. وبوسع التربويين التوجه نحو هذا المصدر لتوفير معلومات موثوقة ومقترحات يمكن استخدامها في موقف مطلوب فيه اتخاذ القرار. وتتوفر هذه الثروة من المعرفة للتربويين نتيجة للبحث العلمي في المشكلات التربوية. ولم تتأثر التربية على الدوام بنتائج استقصاءات منتظمة بعناية كهذه. وفي الواقع، يمكن القول بأن تطور العلم التربوي لا يزال نسبيا في مرحلة مبكرة.

مصادر المعرفة SOURCES OF KNOWLEDGE

وقبل أن نقوم بمتابعة دور الاستقصاء العلمي في التربية، دعونا نتفحص الطرق التي بواسطتها قام الناس على مرّ التاريخ بالبحث عن إجابات لأسئلتهم. ويمكن تقسيم مصادر المعرفة إلى خمسة محاور: 1- الخبرة 2- المرجعية/ السُلطة 3- التفكير الاستنتاجي 4- التفكير الاستقرائي 5- الطريقة العلمية.

الخبرة Experience

إن الخبرة مصدر مألوف لدى الجميع ومستخدم جيدا. فالمرء يتعلم أي طريق يستغرق منه وقتا اقل من غيره أو تقل فيه حركة المركبات الأخرى أو يكون اجمل الطرق بنظره وذلك بعد تجربة المرء لعدة طرق توصله من داره إلى مكان عمله. وبواسطة الخبرة الشخصية يتمكن المرء من إيجاد الأجوبة للكثير من الأسئلة التي تواجهه. إن الكثير من الأمور الحكيمة التي تمر من جيل إلى جيل هي حصيلة للخبرة. وإذا ما كنا غير قادرين على الاستفادة من الخبرة، فإن التقدم سوف يتأخر على نحو بالغ. وفي الواقع، فإن هذه القدرة على التعلمّ من الخبرة تُعتبر ميزة بالغة الأهمية للسلوك الذكي.

ولكن بالرغم من كل فوائدها، فإن للخبرة حدوداً إذا ما اعتُبرت مَصْدرًا للحقيقة. فالكيفية التي يتأثر بها إنسان ما بسبب حادث ما تعتمد على من هو ذلك الإنسان. فسوف يكون لشخصين خبرات بالغة الاختلاف إزاء موقف بعينية. فالغابات ذاتها قد تكون ملاذا لطيفًا

لشخص ما بيد إنها موحشة وتنذر بالخطر بالنسبة إلى شخص آخر. وإذا قام مراقبان بمتابعة ما يجري في قاعة دراسية في الوقت نفسه، فإنهما قد يجمعان تقارير صادقة ومختلفة إذا ما ركز أحدهما على الأمور التي سارت على نحو صحيح، بينما قام الشخص الآخر بالتركيز وكتابة الأمور التي سارت بصورة خاطئة.

وهناك قصور آخر في مسألة الخبرة يتمثل في أن الشخص غالبا ما يحتاج لمعرفة أمور لا يمكن أن يتعلمها الفرد من خلال الخبرة. فالطفل المنطلق على هواه في اكتشاف علم الحساب بمفرده، قد يكون بإمكانه اكتشاف طريقة في الجمع، لكن من غير المحتمل أن يجد وسيلة كفوءة في حساب الجذور التربيعية للإعداد. وربما هناك مدرس في إمكانه من خلال خبرته معرفة عدد التلاميذ في أحد الفصول الدراسية ذات يوم معين، لكنه ليس بوسعه شخصيًا تقدير عدد سكان الولايات المتحدة الأمريكية.

المرجعية / السُلطة Authority

إن تلك الأمور التي من الصعب أو المستحيل معرفتها بواسطة الخبرة الذاتية، تجعل الفرد غالبا ما يلجأ إلى المرجعية/ السُلطة، وهذا معناه أن الشخص يقوم ببحث عن أجوبة لأسئلة من شخص توفرت له الخبرة في معالجة تلك المشكلة، أو كان لديه مصدر آخر من مصادر الخبرة. فنحن نتقبل كلمة أولئك الذين يحوزون على الثقة كمرجعيات، بأنها صادقة. ولمعرفة عدد سكان الولايات المتحدة الأمريكية يحتاج المرء تقارير مكتب إحصاء سكان الولايات المتحدة. والطالب يبحث في القاموس عن اللفظ الصحيح لكلمة ما جاءت أمامه. والناظر يحتاج للتشاور مع أحد المحامين بخصوص مشكلة قانونية في المدرسة. أما المعلم المبتدئ فإنه يحتاج لسؤال معلم ذي خبرة بشأن أية آراء أو مقترحات. وقد يحاول المعلم الجديد تجربة طريقة معينة في تدريس مادة القراءة لأن المشرف قد اقترح جدواها وفائدتها.

ومن خلال التاريخ، يمكن للمرء أن يجد أمثلة تشير إلى الاعتماد على المرجعية/ السُلطة من أجل الحقيقة، وخصوصا خلال القرون الوسطى عندما كان العلماء الأقدمون مثل أفلاطون وأرسطو، وقساوسة الكنيسة الأولون يفضلون على غيرهم كمصادر للمعرفة الحقة- حتى لو كان ذلك على حساب الملاحظة المباشرة أو الخبرة. وبالرغم من أن المرجعيات هي أحد مصادرنا المفيدة جدا في المعرفة، يتوجب على المرء دائما أن يطرح هذا السؤال: "كيف تعلم المرجعية بذلك؟". وسابقًا كان يُفتَرض أن المرجعية على حق بسبب منصبها، مثل منصب الملك، أو الرئيس، أو الكاهن الأعلى. أما اليوم فأننا لا نرغب في الاعتماد على فرد ما كمرجعية لمجرد المكانة أو المرتبة. فنحن نميل لان نتقبل تأكيدات مرجعية ما شريطة أن تعتمد تلك المرجعية في تأكيداتها على الخبرة أو على مصادر معتبرة أخرى من المعرفة.

أما التقاليد والعُرف فإنها تنتمي إلى حد بعيد إلى المرجعيات التي نعتمد عليها من أجل

توفير أجوبة للعديد من الأسئلة المتصلة بمشكلاتنا المهنية واليومية. وبكلمات أخرى، غالبًا ما يتساءل المرء: كيف كان ذلك يَجْري في الماضي؟ وبعد ذلك يستخدم ذلك المرء الجواب على هذا السؤال ليكون دليله في أعماله وأفعاله. فالعرف والتقاليد كانت على وجه الخصوص مؤثرات بارزة في إطار المدرسة، حيث غالبا ما يعتمد المربون على ممارسات سابقة كدليل يعتمد عليه. وعلى أية حال، فإن أي تفحص لتاريخ التربية يكشف بأن الكثير من التقاليد التي كانت سائدة لسنوات طويلة تبين خطؤها بعدئذ، مما أدى إلى رفضها. فمن التعقل تقييم العادة والعرف بعناية قبل أن يتقبلها المرء كحقيقة.

وكمصدر للحقيقة، فإن للمرجعية قصورها بما يتطلب مراعاة ذلك. ففي المقام الأول، قد تقع المرجعية في الخطأ ولا تدعي المرجعيات العصمة من الأخطاء. وقد يجد المرء بأن أهل المرجعية قد يختلفون فيما بينهم بشأن قضايا محددة، وهذا يشير بأن أقوالهم المرجعية غالباً ما تكون عبارة عن آراء شخصية أكثر من كونها حقيقة.

التفكير الاستنتاجي Deductive Reasoning

ربما كان الفلاسفة الإغريق أول من أسهم في بناء طريقة نظامية لاكتشاف الحقيقة. وقد قدم أرسطو وتابعوه استخدام التفكير الاستنتاجي، الذي يمكن أن يوصف بأنه أحد طرائق التفكير التي يتدرج المرء فيها من العام إلى الخاص المحدد مستخدما قواعد المنطق المعمول بها. وهو نظام لترتيب حقائق معلومة من أجل التوصل إلى إحدى النتائج. ويتم هذا باستخدام الحجج/ البراهين المنطقية. وتتكون أية حجة من عدد من القضايا التي بينها علاقات.

وتكون النتيجة هي العبارة النهائية، والبقية- التي تدعى القضايا- تؤلف دليلا داعماً. إن أحد الأنواع الكبرى للتفكير الاستنتاجي هو (القياس المنطقي). والقياس المنطقي يتألف من قضية كبرى وقضية صغرى، تتبعهما نتيجة. وهذا مثال على القياس المنطقي: [كل الناس فانون] (قضية كبرى)، [المِلكَ إنسان] (قضية صغرى)، بناء على ذلك، فالملك فان (النتيجة).

فإذا ما كانت القضايا صحيحة في التفكير الاستنتاجي، تكون النتيجة بالضرورة صحيحة.

ويتيح التفكير الاستنتاجي للفرد القدرة على تنظيم القضايا في أنماط تقدم برهاناً قاطعاً لصحة النتيجة. فالمعجبون بالروايات الغامضة سوف يتذكرون المحقق (شرلوك هولمز) الذي غالبا ما كان يردد: "إنني استنتج..." وذلك عندما كان يجمع مسبقاً حقائق غير مترابطة مسبقا بطريقة توحي بنتيجة غير متوقعة مسبقاً أيضاً.

وعلى أية حال، فإن للتفكير الاستنتاجي عيوبه: فيتوجب على المرء أن يبتدئ بقضية صحيحة من أجل أن يتوصل إلى نتائج صحيحة. فالنتيجة في القياس المنطقي لا يمكنها مطلقاً أن تتجاوز القضايا. ونظراً إلى أن النتائج الاستنتاجية إسهاب بالضرورة يقوم على معرفة موجودة

مسبقاً، فإن البحث العلمي لا يمكن إجراؤه بالاعتماد على التفكير الاستنتاجي وحده وذلك بسبب الصعوبة المتضمنة في بناء القضية الكلية للعديد من العبارات التي تتعامل مع الظواهر العلمية. ويمكن للتفكير الاستنتاجي تنظيم ما قد جرى مَعرفتُه سابقاً، كما يتمكن من توكيد علاقات جديدة إذ يتدرج فيه المرء من العام إلى الخاص، لكن ذلك غير كاف كمصدر لحقيقة جديدة.

وبالرغم من قصوره فإن التفكير الاستنتاجي مفيد في عملية البحث العلمي. فهو يقدم وسيلة تربط النظرية بالملاحظة. فهو يتيح للباحثين الاستنتاج من النظريات القائمة، الظواهر التي ينبغي ملاحظتها. فالاستنتاجات من النظرية تتيح طرح فرضيات تعدّ جزءاً حيوياً من الاستقصاء العلمي.

التفكير الاستقرائي Inductive Reasoning

تكون نتائج التفكير الاستنتاجي صحيحة فقط إذا ما كانت القضايا التي استندت عليها صحيحة. ولكن كيف سيعلم الفرد بأن القضايا صحيحة؟ ففي العصور الوسطى كانت العقائد المذهبية (الدوجماتية) غالباً ما تقوم مقام القضايا الصحيحة، ونتيجة لذلك كانت النتائج غير الصحيحة هي الحصيلة. وكان فرنسيس بيكون (1561-1626) هو أول من دعا إلى إيجاد طريقة جديدة للمعرفة. ولقد اعتقد بأنه يتوجب على المفكرين أن يحرروا أنفسهم من قبول قضايا تفرضها المرجعيات كحقيقة مطلقة. كما اعتقد فرنسيس بيكون بأنه يتوجب على الباحث أن يصل إلى نتائج عامة على أساس حقائق تم جمعها من خلال الملاحظة المباشرة. ولقد نصح فرنسيس بيكون الباحث عن الحقيقة بأن يراقب الطبيعة مباشرة وان يحرّر العقل من التحيز ومن الأفكار الجاهزة سلفاً، والتي قام بتسميتها (أصناما). ويرى فرنسيس بيكون أن الحصول على المعرفة يتطلب من المرء مراقبة الطبيعة نفسها، وتجميع حقائق خاصة، وصياغة تعميمات مما وجد. وأهمية الملاحظة تبرز في الحكاية التالية المنسوبة إلى فرنسيس بيكون:

"في عام 1432، حدث نزاع خطير بين الكهنة حول عدد الأسنان في فم الحصان. ولمدة ثلاثة عشر يوماً استمر الجدال مستعرا دون توقف، وقد تم استعراض جميع الكتب القديمة وسجلات الأحداث التاريخية ومجالات المعرفة الواسعة المملة بما لم تر المنطقة مطلقاً مثيلا لها من قبل. وفي بداية اليوم الرابع عشر طلب قس شاب ذو مظهر وسيم من رؤسائه العلماء الإذن بالكلام، وفور ذلك – وسط تعجب المتجادلين بما أغضب حكمتهم العميقة – التمس منهم التصرف بدون توتر بطريقة فظّة غير مسبوقة من قبل، وذلك بالنظر داخل الفم المفتوح لأي حصان وإيجاد الجواب لتساؤلاتهم. عند ذلك، استشاطوا بغضب كبير لأنهم شعروا بأن كرامتهم قد أهينت، وفي هيجان عام هائل انهالوا عليه ضرباً وقذفوا به إلى الخارج فوراً، قائلين: بالتأكيد أن الشيطان قد أغوى هذا الكاهن المبتدئ الوقح لطرح أساليب غير مقدسة لم تعرف من قبل لإيجاد الحقيقة خلافاً لتعاليم الآباء. وبعد أيام عديدة وكثيرة من هذا الكفاح المضني جاءت

حمامة السلام وحطت رحالها بين المجتمعين، فقام الجميع وكأنهم رجل واحد بالإعلان أن المشكلة سوف تبقى لغزاً غامضاً أزلياً وذلك بسبب الشحة المحزنة للأدلة التاريخية والدينية المناسبة لتلك المسألة وأمروا بتوثيق الموقف كما هو".

لقد كان الشاب في هذه الحكاية يدعو إلى إيجاد طريقة جديدة للبحث عن الحقيقة: ونعني بذلك أن يجدّ المرء في البحث عن الحقائق وليس مجرد الاعتماد على المرجعيات أو على مجرى التكهنات. وهذا الأمر كان سيصبح بعدئذ المبدأ الأساسي لكل العلوم.

وفي نظام فرنسيس بيكون، كانت الملاحظات تجري على أحداث معينة في صنف ما، وتعمم الاستنتاجات على الصنف كله. وتعرف هذه الطريقة بالتفكير الاستقرائي، وهي عكس العملية المستخدمة في الطريقة الاستنتاجية. والفرق بين التفكيرين الاستنتاجي والاستقرائي ربما يمكن مشاهدته في المثالين التاليين:

الاستنتاجي: كل حيوان ثَدْيي لديه رئتان
كل الأرانب حيوانات ثديية
بناء عليه، كل أرنب يملك رئتين

الاستقرائي: كل أرنب قد خضع للمراقبة لديه رئتان
بناء على ذلك، كل أرنب يملك رئتين

على القارئ أن يلاحظ بأنه ينبغي للقضايا أن تكون معروفة في التفكير الاستنتاجي قبل أي نتيجة يتم التوصل إليها، لكن في التفكير الاستقرائي يتم التوصل إلى النتيجة بواسطة ملاحظة الأمثلة، ومن ثم التعميم من الأمثلة إلى الصنف بأكمله. ومن أجل التأكد التام بشأن نتيجة استقرائية، فإنه ينبغي على الباحث أن يُلاحظ جميع الأمثلة. وبموجب نظام فرنسيس بيكون فإن هذا يعرف باسم (الاستقراء التام)، وهذا يتطلب من الباحث أن يفحص كل مثال للظاهرة. وفي المثال أعلاه ومن أجل التأكد التام بأن كل أرنب يملك رئتين يتوجب على الباحث أن يطبق الملاحظات على جميع الأرانب الحية الآن، وكذلك على الأرانب الماضية والمستقبلية. وعمليا، فإن هذا عادة غير ممكن التطبيق. بناء على ذلك، يجب على الفرد أن يعتمد على الاستقراء غير التام القائم على الملاحظة الناقصة.

وتكون النتائج الاستقرائية مطلقة فقط عندما تكون المجموعة التي جرى التأكيد عليها صغيرة. مثلا، ربما يلاحظ شخص ما بأن جميع الطلبة ذوي الشعر الأحمر في فصل معين يحصلون على درجات فوق المتوسطة في التهجئة، ويؤكد منطقيا بأن الطلبة ذوي الشعر الأحمر المسجلين في الفصل يحوزون على درجات أعلى من المعدل في التهجئة. لكن المرء لا يمكنه استخلاص نتائج منطقية بخصوص درجات التهجئة للتلاميذ ذوي الشعر الأحمر في فصول أخرى أو في فصول مستقبلية.

ونظراً لأن بوسع المرء إجراء استقراءات كاملة فقط في حالة المجموعات الصغيرة العدد، فنحن عموما نستخدم الاستقراء الناقص- وهو نظام يقوم فيه شخص ما بملاحظة إحدى العينات لمجموعة ما ومن ثم يستدل من العينة على الخاصية للمجموعة كلها. وكمثال على نتيجة تستند إلى استقراء ناقص يتمثل بالتفكير الراهن حول الخصائص البدنية للأطفال الأذكياء جداً. فلسنوات عديدة مضت، ساد الاعتقاد بصورة عامة بأن الأطفال الأذكياء جدا عرضة لحالة الضعف البدني. وحتى في وقتنا الحاضر، فإن رسامي الصور المتحركة (الكاريكاتير) عادة ما يصورون التلميذ الذكي كمخلوق مهزول وذي نظارات طبية سميكة. أن أحد الرواد في حقل الاختبارات الذهنية- ويدعى تيرمان- كان مهتما بخصائص اليافعين من الأذكياء بصورة استثنائية (1926). وقد قام بدراسة مكثفة لأكثر من ألف طفل في ولاية كاليفورنيا من الذين حصلوا على أكثر من مائة وأربعين درجة في اختبار الذكاء (ستانفرد- بينيه). ولقد وجد هذا الباحث بأن الطول، والوزن، والوضع الصحي العام لأولئك الأطفال فوق المستوى المتوسط بقليل مقارنة مع أطفال آخرين في نفس أعمارهم. ومن تلك الدراسة جاء الاستنتاج بأن الأطفال الأذكياء ليسوا مهزولي البنية حسب التوقع التقليدي إذ كانوا أعلى من المتوسط قليلا في تطورهم البدني من أطفال آخرين يملكون مستويات ذكاء متوسطة.

ويجب على المرء أن يلاحظ بأن هذا الاستنتاج لم تتم البرهنة عليه بصورة قاطعة. وببساطة فإن من المحتمل جداً أن يكون ذلك الاستنتاج صحيحاً. ومن أجل التأكد بصورة تخلو من الشك لهذا الاستنتاج، يتوجب على المرء أن يجري قياسات بدنية لجميع الأطفال الذين حصلوا على مائة وأربعين نقطة أو فوق على مقياس اختبار الذكاء (ستانفرد- بينيه). وحتى في ذلك الحين يمكن للفرد أن يكون متأكداً فقط حول مميزات هكذا أطفال حاليا، ولا يمكن لذلك الفرد أن يكون مائة في المائة متأكدا من أن الشيء ذاته سوف يكون صحيحا ومنطبقا على هكذا أطفال مستقبلا.

وبالرغم من أن الاستقراء الناقص لا يقود إلى نتائج معصومة عن الخطأ، إلا أن في إمكانه إعطاء معلومات يمكن الاعتماد عليها والتي في ضوئها يمكن للمرء اتخاذ قرارات معقولة.

الطريقة العلمية The Scientific Approach

إن الاستخدام المقتصر على الاستقراء غالباً ما نتج عنه تراكم لمعرفة منعزلة ومعلومات كان إسهامها قليلاً في تطور المعرفة. وبالإضافة إلى هذا، فلقد اتضح بأن الكثير من المشكلات لم يكن بالوسع حلها بواسطة الاستقراء وحده. فلقد كان محتما أن يتعلم الباحثون تكامل الأوجه المهمة جدا للطريقتين الاستقرائية - الاستنتاجية في تقنية جديدة، ونعني بذلك: الطريقة الاستقرائية الاستنتاجية، أو الطريقة العلمية. ففي تطويره لنظرية التطور والنشوء، يعتبر تشارلس داروين أول من قام بتطبيق هذه الطريقة في اقتفاء المعرفة. وقد كتب داروين بأنه قد أمضى وقتا طويلا في

القيام بملاحظات بيولوجية، آملا بأن يتمكن من التوصل إلى بعض التعميمات الخاصة بالتطور والارتقاء. وفي القطعة التالية، يصف داروين كيف توصل إلى طريقة جديدة: "إن دفتر ملاحظاتي الأول (حول التطور والنشوء) كان قد ابتدأ في تموز / 1837. فلقد كنت اعمل بناءاً على مبادئ بيكونية (نسبة إلى فرنسيس بيكون)، ودون الاعتماد على نظرية قمت بتجميع حقائق على نطاق واسع، وخصوصا في مجال النتاجات الزراعية والمدجنة، وذلك بواسطة استفسارات مطبوعة والتحدث مع مربين للحيوانات ماهرين بالتحدث إلى الجنائنيين ثم بالقراءة المكثفة. وعندما كنت أشاهد قائمة الكتب المنوعة التي كنت قد قرأتها ولخصتها- وكان ذلك يتضمن مجموعات كاملة من المجلات- كنت اعجب لاجتهادي ومثابرتي. وأدركت بسرعة بأن الانتقاء كان الأساس المهم في نجاح الإنسان في التوصل المفيد إلى أنواع الحيوانات والنباتات. ولكن الكيفية التي يمكن بها تطبيق الانتقاء على كائنات حية تعيش على طبيعتها بقيت لمدة من الزمن إحدى الحالات الغامضة أمامي.

ففي شهر تشرين الأول/ 1838، أي بعد انقضاء خمسة عشر شهراً على بداية قيامي ببحثي المنظم، حدث أن قرأت مقالاً من أجل المتعة وكان عنوانه (مالتوس والسكان). وبسبب كوني مستعداً حينئذ لتقدير الكفاح من أجل البقاء، الذي يأخذ مجراه في كل مكان، من خلال الملاحظة المستمرة والطويلة لعادات الحيوانات والنباتات، فجلب انتباهي فجأة بأنه تحت تلك الظروف نعمل بعض التغيرات المواتية لحفاظ البقاء كما تعمل تغيرات غير مواتية أخرى للفناء. ونتيجة ذلك سوف تنشأ أجناس جديدة. وهكذا حصلت أخيراً على نظرية أستطيع بواسطتها مواصلة عملي. (داروين: صفحة 68، 1899).

إن طريقة داروين التي قامت على الملاحظة فحسب، كانت غير مثمرة، إلى أن قاده قراءة وتفكير أعمق إلى صياغة إحدى الفرضيات التجريبية التي تفيد في تفسير الحقائق التي قام بتجميعها من خلال الملاحظة. بعد ذلك، مضى داروين في اختبار هذه الفرضية من خلال استنتاجات منها ثم قيامه بتجميع بيانات إضافية لتقرير فيما إذا كانت تلك البيانات تؤيد الفرضية أم لا. وبهذه الطريقة في البحث، تمكن داروين من تطوير نظريته في النشوء والارتقاء. وهذا الاستخدام لكل من التفكيرين الاستقرائي والاستنتاجي صفة مميزة للبحث العلمي الحديث الذي يعتبر أفضل طريقة يمكن الاعتماد عليها في الحصول على المعرفة.

وتوصف الطريقة العلمية على أنها عموماً عملية يتحرك فيها الباحثون بصورة استقرائية من ملاحظاتهم إلى تشكيل الفرضيات، ثم يستنتجون من الفرضيات المضامين المنطقية لتلك الفرضيات. ويستدل الباحثون على النتائج التي ستعقب ذلك إذا ما كانت العلاقة التي جرى افتراضها صحيحة. وإذا ما كانت هذه المضامين التي تم استنتاجها تنسجم مع الكيان المنظم

للمعرفة المقبولة آنذاك، فإنه يجري اختبارها ثانية وذلك باستخدام البيانات التجريبية. وعلى أساس الأدلة، فإن الفرضيات يجري قبولها أو رفضها.

إن استخدام الفرضية هو اختلاف أساسي بين الطريقة العلمية والتفكير الاستقرائي. ففي التفكير الاستقرائي يقوم المرء بتسجيل الملاحظات أولا وبعد ذلك يقوم بتنظيم المعلومات التي حصل عليها. أما في الطريقة العلمية، فيقوم المرء بالتفكير فيما عساه أن يجد إذا ما كانت الفرضية صحيحة، وبعدها يسجل ملاحظاته بصورة نظامية من أجل تأكيد أو رفض ما جاءت به الفرضية.

مثـــال على الطريقة العلمية

An Example of The Scientific Approach

يقدم روبرت بيرسيك وصفاً حياً وموجزاً للطريقة العلمية وذلك بمقارنتها مع طريقة صيانة إحدى الدراجات النارية من أجل الإبقاء عليها صالحة للعمل:

"تم استخدام نوعين من المنطق: الاستقرائي والاستنتاجي، فالاستدلالات الاستقرائية تبدأ بملاحظات الآلة للتوصل إلى نتائج عامة. مثال على ذلك، إذا ما ارتطمت الدراجة النارية بشيء صلب في الأرض وحدث خلل ما في الآلة، ثم ارتطمت ثانية بشيء صلب ثان وحدث خلل ما في الآلة، كما حدث ذلك في مرتين ثالثة ورابعة، ثم تحركت فوق أرض ناعمة منبسطة في طريقها ولم يحدث خلل ما في آلتها، عندئذ يتمكن المرء من الاستنتاج بأن الخلل الحاصل في آلة الدراجة كان بسبب تلك الارتطامات. وهذا هو الاستقراء: وهو المحاكمة من خبرات معينة للوصول إلى حقائق عامة.

أما الاستدلالات الاستنتاجية فتعمل عكس ذلك. فهي تبدأ بالمعرفة العامة ثم تتنبأ بملاحظة معنية. مثلا: من القراءة المتسلسلة لحقائق عمل آلة الدراجة، يعلم الشخص الميكانيكي بأن بوق الدراجة يعمل حصرا بالكهرباء القادمة من بطارية الدراجة، بعدئذ يتمكن الميكانيكي منطقيا من الاستنتاج بأنه إذا ما كانت البطارية مستنــزفة فإن البوق لن يعمل. وهذا هو الاستنتاج.

وإيجاد الحل لمشكلات بالغة التعقيد بالنسبة للحس العام، فإن الحل يأتي عن سلسلة طويلة لاستدلالات استقرائية واستنتاجية تنحصر بين آلة الدراجة والتسلسل الهرمي الذهني للآلة، المتوفر في الكتيبات الدليلة. والبرنامج الصحيح لهذا النسيج يدعى رسمياً بالطريقة العلمية.

وفي الواقع لم أشاهد مشكلة صيانة دراجة نارية معقدة بما يكفي في الواقع لاستخدام طريقة علمية شكلية تامة(*). فمشكلات إصلاح الآلات ليست بتلك

(*) شكلي / صوري (formal) مصطلح، يشير إلى الالتزام بالقواعد / المبادئ التي تحدد قالباً منهجياً منطقياً ونظاميا لمعالجة موقف ما. (المراجع)

الصعوبة. وعندما أفكر بالطريقة العلمية الشكلية، ترد إلى عقلي بعض الأحيان صورة ذهنية لتراكتور هائل القوة، بطيء، مضجر، مقعقع، كادح لكنه لا يُقهر. فتأخذ المشكلة ضعف المدة العادية، أو خمسة أضعاف المدة الزمنية، وربما أكثر من ذلك بكثير، لكنك تعلم أنك في النهاية سوف تتمكن منها. وعندما تأتي أمامك إحدى المشاكل العويصة، وتحاول كل شيء في حلها، ثم لا تجد حلا مناسبا لها، فانك حينئذ تدرك بأن الطبيعة قد اتخذت قرارها بأن تكون صعبة، فتقول في نفسك: "حسنا أيتها الطبيعة، هذه نهاية الرجل الطيب".

وعندئذ تستدير نحو الاستعانة بالطريقة العلمية الشكلية. ومن أجل ذلك فانك تحتفظ لنفسك بدفتر ملاحظات مختبري، لتسجل فيه بطريقة منظمة كل ما يلزم، وذلك من أجل أن تعلم في أية وقت من الأوقات أين أنت، وأين كنت، وأين ستذهب، والهدف الذي تسعى إليه. فهذا ضروري في العمل العلمي، وفي تكنولوجيا الإلكترونيات، ودون ذلك فإن المشكلات تصبح معقدة إلى درجة تشعرك بالضياع فيها، وبالحيرة، ونسيان ما كنت تدركه وما كنت لا تدركه، وبالتالي يتحتم عليك التوقف. لكن، في صيانة الدراجة النارية فإن الأمور ليست متشابكة إلى هذا الحد، لكنه عندما يبدو الارتباك، فمن المستحسن السيطرة على الموقف والقيام بتنظيم كل شيء بدقة. وفي بعض الأحيان، فإن مجرد القيام بكتابة المشكلات على الورق يحصر تفكيرك أكثر لفهمها كما هي في الواقع.

فالبيانات المنطقية التي تدخل في دفتر الملاحظات يمكن تصنيفها إلى ستة أقسام: 1- بيان المشكلة، 2- افتراضات بخصوص سبب نشوء المشكلة، 3- تصميم تجارب لاختبار كل فرضية على حدة، 4- نتائج متوقعة للتجارب، 5- النتائج الملحوظة للتجارب، 6- نتائج مستخلصة من نتائج التجارب.

وهذا الأمر ليس مختلفا عن الترتيب الشكلي للكثير من دفاتر الملاحظات المستعملة في الكليات والمدارس الثانوية. لكن الغاية هنا، ليست مجرد الانهماك في العمل. بل الغاية هي التوجه الدقيق للأفكار التي سوف لن يكتب لها النجاح إن لم تكن مضبوطة.

والغاية الحقيقية للطريقة العلمية هي التأكد من أن الطبيعة لم تقم بخداعك في الظن بأنك تعلم شيئا وفي الواقع إنك لا تعلمه. فلا يوجد هناك ميكانيكي أو عالم أو فني على قيد الحياة ممن لم يعان من ذلك كثيراً ولم يكن بغريزته حذراً. وهذا هو السبب العام الذي يجعل الكثير من المعلومات الميكانيكية والعلمية تبدو مضجرة وحذرة. وإذا ما أهملت أو تعاملت رومانسيا بالمعلومات العلمية ولم تعطها استحقاقها، فإن الطبيعة سوف تسخر منك تماما. فالطبيعة تفعل ذلك غالبا وحتى إن لم تتح لها الفرص. لذلك يتوجب على المرء أن يكون منتبها إلى أقصى حد ومنطقيا بصورة صارمة عندما يتعامل مع الطبيعة: فإذا ما كانت هناك زلة منطقية واحدة، انهار الصرح العلمي بأجمعه. وإذا ما كان هناك استنتاج خاطئ واحد بخصوص آلة الدراجة النارية فانك ستجد نفسك معطلا إلى ما لا نهاية.

ففي الجزء الأول من الطريقة العلمية الشكلية- وهي بيان المشكلة- فإن المهارة الأساسية بكل تأكيد، تكمن في بيان ما تعلمه بصورة مؤكدة. فمن الأحسن كتابة جملة: "أعط حلاً للمشكلة: لماذا لا تعمل الدراجة النارية؟" فهذه الجملة تبدو بليدة، لكنها صحيحة. وهذا احسن من كتابة جملة: "أعط حلا للمشكلة: ما هو الخطأ في النظام الكهربائي للدراجة الكهربائية؟" عندما لا تعلم مطلقا أن المشكلة في النظام الكهربائي. وما يتوجب عليك كتابته هو: "أعط حلا للمشكلة: ما هي العلة بالدراجة النارية؟"، وبعد ذلك اكتب في الجزء الثاني من المدخل: "الفرضية الأولى: المشكلة تقع في النظام الكهربائي" فأنت مدعو للتفكير في أكبر عدد من الفرضيات، وبعد ذلك تقوم بتصميم تجارب لاختبارها لترى أياً منها صحيحة وأياً منها خاطئة.

وهذه الطريقة الحريصة للأسئلة الابتدائية تجنبك اتخاذ استدارة كبيرة خاطئة التي ربما تسبب لك أسابيع من الجهد الإضافي، أو ربما توقفك عن العمل تماما. ولذلك فإن الأسئلة العلمية تبدو غبية على السطح، لكن طرحها يجنب الأخطاء الغبية فيما بعد.

والجزء الثالث من الطريقة العلمية الشكلية المدعو بالتجريبية يظنه الرومانسيون أحياناً بأنه العلم بأجمعه إذ أنه الجزء الوحيد الأكثر بروزاً على السطح. فهم يشاهدون الكثير من أنابيب الاختبار والأجهزة الغريبة وأناساً يتراكضون حولها للقيام بالاستكشافات. ولكنهم لا يرون التجربة كجزء من عملية عقلية أكبر، وبذلك فهم غالبا ما يخلطون بين التجارب والعروض التي تبدو متماثلة. فإذا ما قام أحد الأشخاص بعرض علمي قوامه أزيز الصراخ وأجهزة علمية باهظة الثمن تتسبب في هلاك مبدعها، فإنه لا يقوم بأي شيء ذي قيمة علمية إذا ما كان يدري مسبقا ماهية النتائج المترتبة على جهوده تلك. ومن ناحية أخرى، فإن ميكانيكي الدراجة النارية الذي يضغط على مفتاح بوق التزمير ليرى فيما إذا كانت البطارية عاملة أم لا، يقوم بصورة غير رسمية بتجربة علمية حقيقية. إنه يختبر إحدى الفرضيات وذلك بطرحه السؤال على الطبيعة. وعالم التلفاز الذي يتمتم بهدوء: "إن التجربة فاشلة، إذ أخفقنا في الوصول إلى ما كنا نطمح إليه"، يعاني أساساً من كاتب نصوص مسرحية رديء. فالتجربة لا تكون مطلقا فاشلة فقط بسبب إنها تخفق في الوصول إلى نتائج متوقعة. فالتجربة تعتبر فاشلة فقط عندما تخفق في اختبار إحدى الفرضيات المطلوبة، وذلك عندما لا تبرهن البيانات التي تقدمها عن أي شيء.

فالمهارة في هذه المرحلة تتكون من استخدام التجارب التي تختبر الفرضية المطلوبة، لا أكثر من هذا ولا أقل، فإذا ما علا صوت بوق الدراجة النارية، واستنتج الميكانيكي بأن النظام الكهربائي يعمل بكامله، فإن الميكانيكي في ورطة كبيرة. فلقد توصل إلى استنتاج غير منطقي. فصوت البوق يدله فقط بأن البطارية والبوق يعملان. ومن أجل تصميم تجربة جيدة يتوجب على الميكانيكي أن يفكر بصرامة كبيرة بدلالة ماذا يسبب مباشرة ماذا. وهذا الشيء تعرفه من التسلسل المعرفي. فالبوق لا يقوم بتحريك الدراجة. والبطارية أيضاً لا تقوم بذلك إلا بطريقة غير مباشرة جداً. إن النقطة التي تجعل النظام الكهربائي يحرك الدراجة بصورة مباشرة هي شمعات الإشعال، وإذا لم يجر

الاختبار في هذه النقطة للنظام الكهربائي فسوف لن تعرف مطلقا فيما إذا كان الخطأ كهربائيا أم لا.

ومن أجل جودة الاختبار يقوم الميكانيكي بإزاحة شمعة الاشتعال من مكانها ويضع قاعدتها بصورة تلاصقها بالماكنة وذلك من أجل أن تكون متصلة كهربائيا بها، ثم يدور المحرك لثوان قليلة من أجل أن يراقب فتحة شمعة الاشتعال بانتظار شعلة زرقاء اللون. وإذا لم يكن هناك أية شرارة، يمكن للميكانيكي الاستنتاج: (أ) هناك خطأ كهربائي ما، أو (ب) التجربة غير متقنه. وإذا ما كان ذا خبرة، فإنه سوف يحاول مرات أخرى، ثم يقوم بتدقيق التوصيلات، جاهداً بأن يجعل شمعة الاشتعال تعطي شرارتها الزرقاء اللون. فإذا أخفق في مسعاه، فسوف يستنتج بأن (أ) كان الصحيح وأن هناك خطأ كهربائي ما، وإن التجربة قد انتهت. فيكون قد برهن بأن فرضيته صحيحة.

وفي النتائج للتصنيف الأخير، تكون المهارة ببيان ما قادت إليه التجربة ولا أكثر من ذلك. فلم تدل التجربة على أنه عندما يقوم بإصلاح النظام الكهربائي، فإن الدراجة النارية سوف تعمل. فلربما تكون هناك أعطال في أماكن أخرى. لكن الميكانيكي يعلم جيدا بأن الدراجة النارية لن تعمل دون أن يعمل النظام الكهربائي، فيقوم بطرح السؤال الشكلي الآخر: "حل المشكلة: ما هو الخطأ في النظام الكهربائي؟".

عندئذ يطرح الميكانيكي فرضياته حول الأخطاء ثم يقوم باختبارها. فمن خلال الأسئلة الصحيحة، واختيار الاختبارات الصحيحة، ومن ثم الاستنتاجات الصحيحة، يجد الميكانيكي طريقه في النسق المنظم والتسلسلي للدراجة النارية إلى أن يتوصل إلى السبب الدقيق أو الأسباب لعطل الآلة، فيستبدلها لاستبعاد التسبب في عطل الآلة حينئذ.

المراقب غير المتدرب لا يرى إلا العمل المادي، وغالبا ما يفكر في أن العمل المادي هو ما يقوم به الميكانيكي في الأساس. وفي الواقع، فإن العمل المادي هو أصغر وأسهل جزء على الإطلاق مما يقوم به الميكانيكي. وإلى حد بعيد فإن أعظم جزء يقوم به الميكانيكي هو المراقبة الدقيقة والتفكير الصحيح. وهذا يفسر ما يبدو أحياناً على الميكانيكيين من قلة الكلام والانطواء خلال إجراء اختباراتهم. فلا يعجبهم حديثك إذ ينصب تركيزهم على صور ذهنية، وتسلسل النظام، وهم حقيقة لا يوجهون النظر إليك أو إلى بدن الدراجة النارية أبدا. إنهم يستخدمون التجربة كجزء من أحد البرامج لتوسعة هرم معرفتهم بتفاصيل الدراجة النارية العاطلة ثم يقارنونها مع هرم المعلومات الصحيحة في أذهانهم. إنهم ينظرون إلى الشكل الأساسي. (بيرسيك، 1974، ص107-111).

نجد في رواية بيرسيك، خمس خطوات نموذجية في البحث العلمي، وهي:

1- *تشخيص المشكلة:* تتمثل الخطوة الأولى في إدراك وجود مشكلة. وربما تتضمن المشكلة أحد الأسئلة عن شيء ما، أو عن تناقض في التحريات، أو عن فجوة ما في

المعلومات. ففي المثال أعلاه، إن حقيقة أن الدراجة النارية لم تعمل يحدد المشكلة.

2- *بيان أو تقرير المشكلة:* الخطوة التالية تتمثل في توضيح المشكلة، فالباحث يبين بدقة أكثر طبيعة ومدى المشكلة التي تم تشخيصها.

3- *صياغة الفرضيات:* يقوم الباحث بصياغة الفرضيات حول الحلول الممكنة للمشكلة، ففي المثال الحالي، كانت الفرضية الأولى تقول بأن الدراجة النارية لا تبتدئ بالعمل بسبب مشكلة ما في النظام الكهربائي.

4- *توقع النتائج:* بعد ذلك، يقوم الباحث بتوقع النتائج لكل فرضية، ويعني ذلك تحديد النتائج التي يجب ملاحظتها إذا ما كانت الفرضية صحيحة.

5- *اختبار الافتراضات:* يقوم الباحث بجمع البيانات الموضوعية لكي يقيّم كفاية كل فرضية جرت صياغتها. وإذا ما ساندت البيانات الفرضية، فيجري قبولها على أساس أنها تفسير عقلاني مقبول. وإذا لم تدعم البيانات الفرضية، يجري عندئذ رفضها.

طبيعة العلم THE NATURE OF SCIENCE

ربما يمكن، في هذه المرحلة، القول بأن جميع العلوم تشترك في هذه الطريقة العامة من أجل الوصول إلى معرفة يمكن الركون إليها، وذلك بالرغم من اختلاف كل علم عن الآخر في الجوهر أو في تقنياته المميزة. إن هذه هي الطريقة في البحث التي تقرر فيما إذا كان هذا الفرع من المعرفة علما أم لا. ولعل أفضل وصف للعلم يكون بالإشارة إليه كطريقة بحث تسمح للباحثين اختبار الظواهر الطبيعية التي تثير اهتمامهم. وبالإضافة إلى الطريقة التي يتبعها العلماء أثناء بحثهم عن معرفة موثوقة، هناك أوجه أخرى معينة من الطريقة العلمية التي سوف نتفحصها بصورة موجزة. وهذه الأوجه هي: 1- افتراضات أقامها العلماء، 2- مواقف اتخذها العلماء، 3- صياغة النظرية العلمية.

افتراضات أقامها العلماء Assumptions Made by Scientists

إن أحد الافتراضات الأساسية التي أقامها العلماء هي إن الأحداث التي يقومون ببحثها تخضع للقوانين أو منظمة – فلا يوجد هناك أي حدث نزوي. إن العلم يستند إلى الاعتقاد القائل بأن لجميع الظواهر الطبيعية عوامل سبقتها. وهذا الافتراض يدعى أحياناً (الحتمية العامة). ولقد افترض البدائيون أسبابا خارقة للطبيعة لأغلب الأحداث التي لاحظوها. والعلم الحديث لم يتطور إلى أن بدأ الناس يتجاوزون التفسيرات خارقة الطبيعة وبدءوا بالاعتماد على مراقبة الطبيعة ذاتها من أجل إيجاد الأجوبة.

وهذا الافتراض يشكل الأساس لأي بيان يعلن بأن أحداثا معينة ستقع تحت ظروف خاصة. فالكيميائي، مثلاً، قادر على أن يعلن بأنه إذا تم تسخين خليط من كلورات البوتاسيوم وثاني أكسيد المنجنيز فينتج الأكسجين. والعلماء السلوكيون بذات الأسلوب، يفترضون بأن سلوك الكائنات الحية يخضع للقوانين وقابل للتنبؤ.

وثمة اعتقاد، ذو صلة بالافتراض الأول، يرى أن الأحداث في الطبيعة هي مرتبة ومنتظمة إلى درجة ما على أقل تقدير، وأن هذا الترتيب وذلك التنظيم في الطبيعة يمكن اكتشافهما من خلال الطريقة العلمية.

والافتراض الثاني يقول بأنه يمكن استخلاص الحقيقة في النهاية، فقط، من الملاحظة المباشرة، والاعتماد على الملاحظة التجريبية يميز بين ما هو علمي وغير علمي. فالعالم لا يعتمد على مرجعية كمصدر للحقيقة، لكنه يصر على دراسة الدليل التجريبي. وفي تاريخ العلوم، نجد عدة أمثلة لعلماء رفضوا المفاهيم السائدة في عصرهم آنذاك واستمروا في ملاحظاتهم وتجاربهم. فالتجارب المبكرة للعالم (غاليلو) على الأجسام الساقطة من أعلى قادت إلى معرفة جديدة ناقضت المفاهيم التي كانت مرجعيات عصره متمسكة بها.

والنتيجة لهذا الافتراض هو الاعتقاد القائل بأن الظواهر الطبيعية التي يمكن إدراك وجودها فعلياً، فحسب، هي التي تقع في نطاق البحث العلمي.

اتجاهات العلماء Attitudes of Scientists

يدرك العلماء اتجاهات مميزة معينة يكتسبونها أثناء متابعتهم لعلمهم:

1- العلماء في الأساس شكاكون. فهم يحتفظون باتجاه شاك جداً حيال بيانات العلم. ويعتبرون الاكتشافات مؤقتة ولا يقبلها العلماء إلا إذا أمكن التحقق من صحتها. ويحتاج التحقق أن يتمكن الآخرون من إعادة الملاحظات والحصول على نفس النتائج. ويود العلماء اختبار الآراء و الأسئلة الخاصة بالعلاقات بين الظواهر الطبيعية. والأكثر من ذلك، يقوم العلماء بإعلان طرائق اختبارهم للآخرين ليتحققوا أو يخفقوا في التحقق، من استنتاجاتهم.

2- العلماء موضوعيون نزيهون. ففي أثناء قيامهم بالملاحظات وتفسير البيانات لا يسعى العلماء للبرهنة على صحة نقطة ما. فهم يبذلون عناية خاصة بجمع البيانات بطريقة لا تسمح لانحيازاتهم الشخصية التأثير على ملاحظاتهم. فهم ينشدون الحقيقة، ويقبلون الحقائق حتى لو كانت متعارضة مع آرائهم الذاتية. وإذا أخلّت الأدلة المتجمعة بموازين إحدى النظريات المفضلة، عندها يقوم العلماء بنبذ تلك النظرية أو تعديلها لتتفق مع البيانات الحقيقية.

3- يتعامل العلماء مع الحقائق، وليس مع القيم. لا يشير العلماء إلى أية مضامين أخلاقية محتملة لاستنتاجاتهم، كما أنهم لا يصدرون أحكاماً لنا بخصوص ما هو جيد وما هو سيئ.

فالعلماء يقدمون البيانات الخاصة بالعلاقة القائمة بين الأحداث، لكن يتوجب علينا أن نتخطى هذه البيانات العلمية إذا ما كنا ننشد قراراً حول النتيجة ما إذا كانت مرغوبة. وهكذا، بالرغم من أنه قد تكون لاكتشافات العلم أهمية كبرى في حل مشكلة ما تتضمن قراراً قيماً، فإن البيانات ذاتها لا تعطي حكما قيمياً.

4- لا يكتفي العلماء بحقائق معزولة، لكنهم ينشدون دمج وتصنيف نتائج بحوثهم. فهم يريدون توظيف الأمور المعروفة في نظام منهجي. وهكذا يطمح العلماء في نظريات تسعى لدمج البحوث التجريبية في نمط ذي معنى. وعلى أية حال، يعتبر العلماء هذه النظريات تجريبية ومرحلية، وخاضعة إلى إعادة النظر كلما وجد هناك دليل جديد.

النظرية العلمية Scientific Theory

إن الوجه الأخير للطريقة العلمية الذي ينبغي أن ينال الاهتمام هو بناء النظرية. فالهدف النهائي للعلم هو تكوين النظرية. فمن خلال الاستقصاء العلمي يقوم العلماء بجمع حقائق كثيرة. ولكن في الوقت الذي تتجمع فيه هذه الحقائق، تبرز الحاجة إلى التجميع، والترتيب، والتصنيف، لإعطاء معنى لهذه الاكتشافات المنعزلة.

ويتوجب تشخيص العلاقات المهمة في البيانات وتفسيرها. وتجري صياغة النظريات من لإيجاز وتنظيم المعرفة الراهنة في مجال معين. وربما يجري تعريف النظرية على أنها "مجموعة من المفاهيم ذات العلاقات المتبادلة، والتعريفات، والقضايا، التي تطرح نظرة منهجية للظواهر، وذلك بتحديد العلاقات بين المتغيرات، بهدف تفسير الظواهر والتنبؤ بها" (Kerlinger, 1986, p. 9).

وترتبط النظريات الملاحظات مع بعضها، بما يتيح الفرصة للعلماء صياغة عبارات عامة حول المتغيرات والعلاقات بينها. وتمتد النظريات من تعميمات قليلة بسيطة إلى صياغات قوانين معقدة. فمثلاً، يلاحظ أنه إذا حافظ الضغط على حاله، فإن غاز الهيدروجين يتمدد إذا ما ازدادت درجة حرارته من 20 إلى 40 درجة مئوية. كما يلاحظ أنه إذا حافظ الضغط على حالة، فإن غاز الأكسجين ينكمش إذا ما انخفضت حرارته من 60 إلى 50 درجة مئوية. والنظرية المعروفة باسم (قانون تشارلس) توجز التأثيرات الملحوظة عن التبدلات في درجة الحرارة على حجوم الغازات: عندما يثبت الضغط وفي الوقت ذاته تزداد درجة حرارة الغاز، فإن حجمه يزداد، وعندما تنخفض درجة الحرارة، يقل حجم الغاز. وهذه النظرية لا توجز المعلومة السابقة فحسب بل تتنبأ بظواهر أخرى إذ تبين ما الذي نتوقعه من أي غاز ما عندما تتبدل درجة حرارته.

أهداف النظريات Purposes of Theories

تقوم النظريات بعمل مفيد في تطور العلوم. ففي المقام الأول، فإن النظرية ترتب الاستنتاجات من بين الكثير من البحوث المتفرقة وتجمعها في إطار يقوم بتهيئة تفسيرات للظواهر. فالنظرية تبين المتغيرات ذات الصلة، والكيفية التي تربط بينها. وعلى سبيل المثال قد تفسر نظرية حول التعلم، العلاقات بين سرعة وكفاءة التعلم ومتغيرات مثل التحفيز، والثواب، والممارسة، وغيرها.

ويتمكن العلماء من المضي في التنبؤ والتحكم في نهاية الأمر، وذلك من الإطار التفسيري لأية نظرية. وحالما يتم تقديم بيان (نظرية) حول العلاقة بين البعوضة المسماة (أنوفيلس) والملاريا في بني البشر، عندئذ يتمكن العلماء من 1- تفسير أسباب استيطان الملاريا في بعض المناطق وليست في أخرى، 2- التنبؤ بالكيفية التي تصاحب فيها التحولات في البيئة تلك التحولات في حدوث الملاريا. 3- السيطرة على الملاريا بواسطة إحداث تحولات في البيئة.

وتثير النظريات عملية تطور المعرفة. فالاستنتاجات من أية نظرية تسمح بالتنبؤات بحدوث الظواهر حيث بعضها لم تجر ملاحظته بعد. مثلاً، تنبأ علماء الفلك، من النظرية، بوجود كواكب بعيدة جداً قبل ملاحظتها فعليا.

واختبار الاستنتاجات من أية نظرية يؤكد النظرية ويطورها. وإذا لم تدعم نتائج البحث، النظرية، يقوم العالم بتنقيحها وبعد ذلك يجمعون بيانات أكثر لاختبار النظرية المنقحة.

الخصائص المميزة للنظريات Characteristics of Theories

إذا ما كان لنظرية أن تخدم هدفها في مجال العلم، فلابد لها أن تخضع لمعايير معينة. وهذه بعض الخصائص التي تميز أية نظرية منطقية:

1- يجب أن تكون النظرية قادرة على توضيح الحقائق الملحوظة ذات الصلة بمشكلة معينة، فيجب أن يكون بوسع النظرية اقتراح السبب المتعلق بالظاهرة ذات العلاقة. ويجب أن يكون هذا التوضيح للأحداث في أبسط شكل ممكن. فالنظرية التي لها تعقيدات وافتراضات أقل يتم تفضيلها على أخرى أكثر تعقيداً. وهذا الشرح يسمى (مبدأ الاقتصاد).

2- يجب أن تنسجم النظرية مع الحقائق الملحوظة، ومع الكيان الراسخ للمعرفة السابقة، فنحن نبحث عن النظرية التي تقدم الطريقة الأكثر احتمالا، أو الأكفأ على الإطلاق في تعليل الحقائق المتراكمة.

3- يجب على النظرية أن تقدم وسائل تحقيقها. ويتم هذا الأمر مع أغلب النظريات وذلك بوضع استنتاجات على شكل فرضيات تبين نتائج يتوقع المرء ملاحظتها إذا ما كانت النظرية صحيحة. عندئذ يكون العالم قادرا على بحث أو اختبار هذه الفرضيات بصورة تجريبية من أجل أن يقرر فيما إذا كانت البيانات تساند النظرية أم لا. ويجب هنا التأكيد على أنه ليس من

المناسب الحديث عن صحة أو خطأ أية نظرية. إن قبول أو رفض أية نظرية يعتمد في الأساس على مدى فائدتها. فالنظرية تكون مفيدة أو غير مفيدة، وذلك يعتمد على مدى الكفاءة التي تؤدي إلى تنبؤات تخص نتائج ملحوظة يتم التأكد منها بعد تجميع البيانات التجريبية. ومع ذلك، فإن أية نظرية تعتبر مؤقتة وخاضعة للتنقيح حينما تتراكم أدلة جديدة.

4- يجب على أية نظرية أن تثير اكتشافات جديدة، وأن تشير، علاوة على ذلك، إلى مجالات في حاجة للاكتشاف.

إن هدف تكوين النظرية قد تم التوصل إليه إلى مدى أبعد في العلوم الطبيعية مما هو عليه في العلوم الاجتماعية، وهذا شيء غير مدهش إذ إنها علوم أقدم. وفي الأيام المبكرة لأي علم من العلوم، فإن التأكيد بصورة نموذجية يكون على التجريبية، حيث يوجه العلماء اهتمامهم إلى جمع الحقائق في مجالات مشكلات معينة. ومع النضوج فقط يبدأ العلم بضم المعارف المعزولة ودمجها في إطار نظري.

وقد عانت التربية بوجه خاص من غياب التوجهات النظرية، حيث انصب التأكيد الأساسي على التجريبية. وجاء انتقاد التربويين لاهتمامهم المستمر (بالحصول على الحقائق) أكثر من (البحث عن الأسباب). ويعكس هذا الاهتمام الأعداد الضخمة للحقائق التي تراكمت خلال الدراسات التربوية، ولكن دون تكامل مصاحب في النظريات من أجل تفسير الظواهر التربوية. والتربية في حاجة إلى تركيز أكثر على تطوير النظرية من أجل الحصول على إدراك أوسع للمشكلات التربوية، وكذلك من أجل إرشاد جهودها على المستوى التجريبي.

وبالرغم من وجود اختلافات واضحة في عدد وقوة النظريات التي تم توطيدها في العلوم الطبيعية والاجتماعية، إلا أن للنظرية الدور ذاته الذي تلعبه في تقدم أي علم. وبغض النظر عن موضوع الدراسة، فإن النظرية تعمل بنفس الطريقة. فالنظرية تقوم باختصار المعرفة القائمة، وتفسير الأحداث الملحوظة والعلاقات، والتنبؤ بوقوع الحوادث غير الملحوظة والعلاقات. ويمكننا القول بأن النظريات تمثل أفضل جهودنا على الإطلاق في فهم البنية الأساسية للعالم الذي نعيش فيه.

قصور الطريقة العلمية في العلوم الاجتماعية

Limitations of The scientific Approach in the social Sciences

لم تحظ التربية والعلوم الاجتماعية الأخرى بالمكانة العلمية النموذجية للعلوم الطبيعية بالرغم من استخدامها للطريقة العلمية وتجميع كمية كبيرة من المعرفة الموثوقة. فالعلوم الاجتماعية لم تستطع بناء تعميمات مكافئة لنظريات العلوم الطبيعية في مدى قوتها التفسيرية، أو في المقدرة على تقديم تنبؤات دقيقة. وهنالك غالبا افتقار إلى الاتفاق بين الباحثين في العلوم الاجتماعية حول الحقائق المعترف بها، أو حول التفسيرات المقبولة للحقائق المسلم بصحتها. وربما لن ترى العلوم الاجتماعية أهداف العلم على الإطلاق تماما مثل ما فعلته العلوم الطبيعية. وهنا يجب

التأكيد على أن الطريقة العلمية في حد ذاتها ليست شرطاً كافياً للتحصيل العلمي. فهناك العديد من العيوب موجودة ضمن تطبيق الطريقة العلمية في التربية وفي العلوم الاجتماعية الأخرى.

تعقيد موضوع البحث Complexity of Subject Matter

إن التعقيد المتأصل لموضوع البحث في العلوم الاجتماعية يشكل عائقاً كبيراً. فالعلماء الطبيعيون عادة يتعاملون مع الظواهر الطبيعية والبيولوجية. وهناك عدد محدود من المتغيرات التي يمكن قياسها بدقة تستخدم في تفسير الكثير من هذه الظواهر، كما يكون محتملاً التوصل إلى قوانين عامة. فمثلا، إن قانون (بويل) حول تأثير الضغط على حجم الغازات- يتعامل مع متغيرات غير معقدة نسبيا – يقوم بصياغة علاقات بين الظواهر تتمتع ظاهراً بالثبات في جميع أنحاء العالم.

ومن الجهة الأخرى، فإن العلماء الاجتماع يتعاملون مع الموضوع الإنساني. فهم يهتمون بسلوك وتطور الإنسان بكونه فرداً وعضواً في إحدى الجماعات. وهناك عدة متغيرات تعمل بصورة مستقلة وأخرى متفاعلة تبادلياً ينبغي أخذها في الاعتبار في أية محاولة لفهم السلوك المعقد للإنسان. فكل إنسان فريد في الطريقة التي يتطور بها، وفي القدرة الذهنية، وفي السلوك العاطفي والاجتماعي، وفي الشخصية الشاملة. وسلوك الناس داخل الجماعات، وتأثير سلوك أعضاء الجماعة على الفرد، يوجب على علماء الاجتماع معالجتها. فمجموعة من تلامذة السنة الأولى لن تسلك في أحد المواقف مثل تلامذة السنة الأولى في موقف آخر. فهنالك متعلمون، ومعلمون، وبيئات، ولكل منها متغيرات تساهم في الظواهر السلوكية الملحوظة في وضع ما. وهكذا يتوجب على الباحثين أن يكونوا حذرين جداً في إصدار التعميمات وذلك لأن البيانات الخاصة بإحدى المجموعات – في موقف ما – قد تتمتع بصدق محدود لدى مجموعات أخرى في أوضاع أخرى.

صعوبات الملاحظة(*) Difficulties in Observation

الملاحظة أمر لابد منه في العلم، وهي أكثر صعوبة في العلوم الاجتماعية مقارنة مع العلوم الطبيعية. فالملاحظة في العلوم الاجتماعية أكثر ذاتية لأنها غالباً ما تتضمن تفسير الملاحظين. ومثال ذلك، حال موضوع البحث الذي غالباً ما يكون استجابات ردود فعل شخص ما إزاء سلوك الآخرين. فالدوافع، والقيم، والاتجاهات ليست جلية للفحص الدقيق، ويتحتم على الملاحظين إعطاء تفسيرات ذاتية عندما يقررون بأن السلوكيات التي لاحظوها تشير إلى وجود دافع، أو قيمة، أو اتجاه ما. وتكمن المشكلة في أن القيم والاتجاهات الذاتية لعلماء الاجتماع قد تؤثر على ما يختارونه للملاحظة، وكذلك على تقديرهم للاستنتاجات التي يسندون إليها نتائجهم. أما العلماء الطبيعيون فإنهم يدرسون الظواهر التي تحتاج إلى تفسير أقل ذاتية.

(*) الملاحظة/ المشاهدة: هي عملية توجيه الانتباه الدقيق نحو موقف ما لفهمه، وتستخدم كأداة في العلوم التجريبية لجمع البيانات (المراجع).

صعوبات في التكرار/ الإعادة Difficulties in Replication

يتمكن الكيميائي بصورة موضوعية من ملاحظة التفاعل بين مادتين كيميائيتين في أنبوبة اختبار. ويمكن نشر الاستنتاجات، كما يمكن للملاحظات أن تتكرر من أشخاص آخرين. أما التكرار في العلوم الاجتماعية فإنه أمر أبلغ صعوبة. إن أي معلم أمريكي لا يتمكن من تكرار الشروط لطريقة تدريس تجريبية لمعلم روسي بنفس دقة قيام أحد الكيميائيين الأمريكيين في إعادة تجربة أحد الكيميائيين الروس. وحتى داخل إحدى بنايات مدرسة واحدة لا يتمكن شخص ما من إعادة أحد المواقف الناشئة بدقة وشمولية. فالظواهر الطبيعية أحداث فريدة لا يمكن إعادتها من أجل غايات الملاحظات.

التفاعل بين الملاحظ الأفراد (*) Interaction of Observer and Subjects

وهناك مشكلة إضافية تتمثل في أن مجرد ملاحظة الظواهر الاجتماعية قد ينتج عنها تغيرات قد يتعذر وقوعها بطريقة أخرى. فلربما يظن الباحثون أن (X) تسبب (Y)، بينما، في الواقع، قد تكون ملاحظاتهم لـ (X) تتسبب في (Y). ومثال ذلك، تجارب (هوثورن) المشهورة التي وجدت أن التغيرات في إنتاجية العمال تعود إلى معرفة العمال بأنه جرى استهدافهم للاستقصاء. فالباحثون من البشر، ولربما يتسبب وجودهم كملاحظين، في موقف، في إحداث تغيير في سلوك العناصر البشرية للبحث. واستخدام آلات تصوير وأجهزة تسجيل خفية قد يساعد في تقليص هذا التفاعل في بعض الحالات، لكن الكثير من البحوث في العلوم الاجتماعية يتضمن استجابات العناصر الإنسانية للبحث لملاحظين من بني البشر.

الصعوبات في الضبط/ التحكم Difficulties in Control

إن مجال الاحتمالات الممكنة للتجارب الخاضعة للتحكم في العناصر الإنسانية محدود أكثر مما هو عليه في العلوم الطبيعية. والتعقيدات المتضمنة في البحث الجاري على العناصر الإنسانية تطرح مشاكل في التحكم لا مثيل لها في العلوم الطبيعية. فالتحكم الصارم في الأخيرة بالشروط التجريبية أمر ممكن داخل المختبر. وسيطرة كهذه غير ممكنة الحدوث مع عناصر بشرية، فينبغي على علماء الاجتماع أن يتعاملوا مع متغيرات كثيرة العدد في آن واحد، ويتحتم عليهم أن يعملوا تحت ظروف أقل دقة. ويحاول هؤلاء العلماء تحديد أكبر عدد ممكن من هذه المتغيرات والتحكم فيها على قدر الإمكان، لكن المهمة في بعض الأحيان صعبة جدا.

مشكلات القياس Problems of Measurement

يتطلب القيام بالتجارب إجراء القياسات للعوامل المشاركة في التجربة. وأدوات القياس في العلوم الاجتماعية أقل بكثير في دقتها وكمالها من أدوات العلوم الطبيعية. وليس لدينا ما يمكن

(*) تشير كلمة "Subject" إلى أحد عناصر ، أفراد عينة البحث العلمي (المراجع).

أن نقارن به في دقة المسطرة، أو ميزان الحرارة، أو بقية أدوات المختبر العديدة. ولقد كنا قد أشرنا آنفا بأن فهم السلوك الإنساني يتعقد لكثرة المتغيرات الفاعلة التي يجب تحديدها، أكانت مستقبلة أو متفاعلة معاً. فالوسائل الإحصائية متعددة المتغيرات، المتوفرة لتحليل الإحصائيات في العلوم الاجتماعية تولي اهتماماً بعدد قليل نسبياً من العوامل المتفاعلة على نحو واضح. والأكثر من ذلك، فإن هذه الوسائل تسمح للشخص بأن يعزو التباين إلى عوامل تعمل أثناء القياس فحسب. أما العوامل التي أثرت في التطور في الماضي فلا يمكن قياسها في الحال الراهن، ولو أنها قد قامت بتأثير مهم في مجرى التطور.

ونظراً لأن البحث في العلوم الاجتماعية معقد بسبب تعقيد وصعوبة الملاحظة، والتكرار والقياس، يتوجب على الباحثين ممارسة الحذر في إصدار التعميمات من دراساتهم. وغالبا ما يكون ضروريا القيام بعدة دراسات في مجال ما قبل محاولة صياغة التعميمات. وإذا ما كانت الاستنتاجات الأولية مؤكدة بانتظام، عندئذ تكون هناك ثقة أكبر في تكوين تعميمات واسعة.

وبالرغم من العوائق، فإن التربية والعلوم الاجتماعية تقدمت كثيراً، ومن المتوقع أن تزداد مكانتها العلمية عندما يصبح الاستقصاء العلمي ومنهجيته أكثر صرامة وتنظيماً.

طبيعة البحث THE NATURE OF RESEARCH

إن البحث العلمي هو تطبيق للطريقة العلمية في دراسة مشكلة ما. وهو وسيلة للحصول على معلومات مفيدة يمكن الاعتماد عليها. وغايته اكتشاف أجوبة لأسئلة ذات معنى وذلك من خلال تطبيق إجراءات علمية. ويجب أن يتضمن أي استقصاء الطريقة التي وصفناها في الفقرة السابقة لكي يجري تصنيفها كبحث علمي. وبالرغم من أن البحث العلمي يجري في أوضاع (مكانية وزمانية) مختلفة كما يستفيد من طرائق مختلفة، فإنه، على صعيد عام، نظامي وبحث موضوعي للوصول إلى معرفة موثوقة.

البحث التربوي Educational Research

عندما تطبق الطريقة العلمية لدراسة مشكلات تربوية، فالحصيلة هي البحث التربوي. فالبحث التربوي هو الطريقة التي يحصل بها الفرد على معلومات معتمدة ومفيدة تخص العملية التربوية، وهدفها هو اكتشاف مبادئ عامة أو تفسيرات للسلوك يمكن استخدامها في الشرح، والتنبؤ، والتحكم بما يتعلق بأحداث في مواقف تربوية – وبكلمات أخرى، في صياغة نظرية علمية.

إن قبول الطريقة العلمية في مجال التربية وفي العلوم الاجتماعية الأخرى قد تختلف كثيراً، مقارنة بقبولها في العلوم الطبيعية. وفي سنة 1897، وجد (رايس) أحد الرواد في مجال البحث التربوي أنه في وضع مشابه لذلك المشهد الذي يُعزى وصفه لفرانسيس بيكون في هذا الفصل.

وقد سأل (رايس) التربويين في اجتماعهم السنوي في قسم الإشراف التربوي عما إذا كان بالوسع تقرير فيما إذا كان الطلبة الذين يمنحون أربعين دقيقة لتعلم التهجئة يوميا أفضل في مستواهم اللغوي من الطلبة الذين يمنحون عشرة دقائق يوميا. ولقد نقل (رايس) ما يلي:

"... لشدة عجبي، وجدت بأن هذا السؤال قد ألقى الفزع في وسط المعسكر. وأول من استجاب كان أحد الأساتذة المعروفين في علم النفس والذي كان منشغلاً في تدريب أساتذة في غرب أمريكا. فقد قال بأنه تتعذر الإجابة على هذا السؤال على نحو مطلق. وقد قام بتقريعي بقسوة لاشغالي الوقت الثمين لعدد مهم من التربويين بطرح أسئلة غبية ". (Rice, 1912, pp. 17-18).

وفي الحقيقة قام (رايس) بتجميع أدلة تجريبية على استفساره، حيث وجد بأن الاختلافات في التحصيل بين الذين يقضون عشر دقائق يوميا وأولئك الذين يقضون أربعين دقيقة يوميا كانت مهملة. كما أشار كذلك إلى أن كثيراً من الكلمات التي تحتم على التلاميذ تعلمها كان ذات أهمية عملية قليلة. وقد قاد عجلة باحثين آخرين- مثل إدوارد ثورندايك- لاستخدام تحليل موثق لتحديد تكرار استخدام الكلمات في لغتنا. وقد أدى عملهم بدوره، إلى تحسينات في العرض اللغوي للكتب وفي المناهج.

وبالرغم من حداثة البحث التربوي، إلا أنه استمر بالتقدم منذ بداياته في أواخر القرن التاسع عشر. وبكونه علماً، فإن البحث التربوي يستخدم طرائق بحثية تتماشى مع المبادئ الأساسية ومع المفاهيم المطبقة في العلم. وهكذا فإن البحث يتضمن عددا من المراحل التي سنقوم بدراستها بعد قليل. وقبل كل شئ، سنلقي نظرة عامة موجزة على نماذج البحث التربوي.

أصناف البحث التربوي Types of Educational Research

إن الأصناف الأساسية للبحث التربوي هي اثنان: البحث النوعي، والبحث الكمي. (أما الصنف الثالث- وهو البحث التاريخي- فيقوم بتجميع بيانات من أجل تكوين الحقائق والوصول إلى نتائج تخص أحد الأحداث الماضية. وسوف نتفحص البحث التاريخي في الفصل الثالث عشر). والبحث الكمي يستخدم قياساً موضوعيا وتحليلات عددية للبيانات للسعي إلى توضيح التغيرات في الظواهر الاجتماعية. ويبدأ هذا النوع من البحث عادة بفرضيات تؤكدها أو تنفيها البيانات. أما البحث النوعي فإنه يسعى إلى (فهم كامل) للظواهر الاجتماعية من خلال الانغماس الشامل للباحث في الموقف. والبحث النوعي عادة يبدأ بالفرضيات، بالرغم من أن البحث قد يولدها حسبما تقع الأمور. وربما يقال بأن البحث الكمي يسعى إلى التفسير، بينما يولي البحث النوعي اهتماما أكثر بالفهم.

وكل من المنهجين الكمي والنوعي، ذو قيمة للباحث التربوي. وتعتمد الطريقة التي يختارها الباحثون على طبيعة السؤال الذي يطرحونه.

البحث الكمي Quantitative Research

يمكن للبحث الكمي أن يصنف أيضاً إلى تجريبي، أو غير تجريبي.

البحث الكمي التجريبي: يقوم الباحث التجريبي بإدخال معالجات(*) مختلفة (تسمى متغيرات مستقلة) إلى مجموعتين أو أكثر، وبعد ذلك يقارن بين المجموعات حتى يحدد تأثير المعالجة على متغير آخر (يسمى المتغير التابع). افرض، مثلاً، أن أحد باحثي الجامعة أراد أن يستقصي تأثير تقديم تغذية راجعة لطلبة، باستخدام الحاسوب، فوراً عقب امتحانات المقرر.

فيقوم الباحث باختيار شعبتين من طلبة الاقتصاد الذين يعلمهم أستاذ واحد، ثم يختار الباحث بصورة عشوائية إحدى الشعبتين لتلقي تغذية راجعة عن طريق الحاسوب حول أداء الطلبة في أسئلة الاختبار. أما الشعبة الأخرى فتتلقى تغذية راجعة بالطريقة التقليدية في حجرة الدراسة. ثم يقوم الباحث بمقارنة الشعبتين معتمداً على درجات اختبار الطلبة والدرجات النهائية التي نالها الطلبة في المقرر الدراسي (متغيرات تابعة).

فإذا ما كانت درجات الاختبار والدرجات النهائية أعلى بكثير في الشعبة التي تلقت تغذية راجعة باستخدام الحاسوب، فيمكن للباحث أن يستنتج بصورة تجريبية بوجود دليل يفيد بأن التغذية الراجعة عن طريق جهاز الحاسوب (المعالجة) قد ساهمت في زيادة التعلم.

وأثناء التجارب يحاول المرء أن يتحكم بجميع المتغيرات التي قد تؤثر على المتغير التابع. وفي المثال السابق، قد يحاول الباحث التأكد من أن كلتا المجموعتين درست على الأستاذ ذاته، وإن كلتا المجموعتين التقت في نفس الوقت من اليوم ذاته، وإن ملحوظات المحاضرة، والقراءة والاختبارات، كانت جميعا نفسها للمجموعتين، وهكذا. وربما يتفحص الباحث مستوى القدرة وخلفية الطلبة في كلتا الشعبتين للتأكد من أن إحدى المجموعتين لم تكن متفوقة على الأخرى. وسوف نناقش البحث التجريبي بدقة أكبر في الفصل التاسع.

البحث الكمي غير التجريبي: في البحث الكمي غير التجريبي، يحدد الباحث المتغيرات، وقد يتحرى عن العلاقات بينها، لكن الباحث يفعّل(**) تلك المتغيرات. أما الأنماط الرئيسية للبحوث غير التجريبية فهي العلية المقارنة والارتباطية والمسحية.

والبحث العلي المقارن يشبه التجربة باستثناء قيام الباحث بتفعيل المتغير المستقل، إذ أن التفعيل قد حدث فعلا أثناء المجرى الطبيعي للأحداث. بعد ذلك، يقارن الباحث مجموعات

(*) المعالجة (treatment): مصطلح، يشير إلى المتغير التجريبي (متغير مستقل) يستخدمه الباحث للتأثير على مجموعة المعالجة/ التجريبية لدراسة الفروق بين المجموعة التجريبية والمجموعة الضابطة (control group) التي لم تتعرض لمتغير المعالجة، وذلك فيما يخص موقفا ما (المراجع).

(**) يفعّل (manipulate)– بتشديد العين – نراها أفضل ترجمة لهذه الكلمة إذ تشير إلى القيام بالتأثير على المتغير المستقل، وهناك من يترجمها (يتلاعب) (المرجع).

تختلف تأثيرات المتغير المستقل فيها وذلك لتحديد التأثير على المتغير التابع.

فمثلاً، للإجابة على هذا السؤال: (هل يواجه أطفال الأول الابتدائي، الذين يعيشون مع واحد من الأبوين فقط، صعوبة في المدرسة أكثر من أقرانهم الذين يعيشون مع الأبوين؟) هناك حاجة لاستخدام الطريقة العلّية المقارنة، ولن يتمكن الباحث من التلاعب بخلفيات عائلات الأطفال بجعل إحداها تعود لأحد الأبوين فقط أو للأبوين معاً، بل عوضا عن ذلك، يقوم الباحث بتحديد الأطفال ذوي الأب أو الأم بمفردهما ويقارن أداء أولئك الأطفال مع أقرانهم من ذوي الأم والأب معا.

وبسبب العجز عن التحكم بعوامل عديدة، يجب على المرء أن يكون حذراً خصوصا في تفسير النتائج للدراسات العلّية المقارنة..

وسوف نلقي نظرة على العلّي المقارن في الفصل العاشر.

أما البحوث الارتباطية فإنها تتوخى تحديد العلاقات بين أثنين أو أكثر من المتغيرات. ويجري التعبير عن مدى العلاقة بشكل مؤشر عددي. وكأمثلة على أسئلة البحث الارتباطي هي: هل هناك علاقة بين نوعية نماذج كتابة طلبة كلية جدد وبين أدائهم خلال السنة الأولى في الكلية؟ ما هي العلاقة بين متغيرات غير معرفية معينة مثل مفهوم الذات الأكاديمي، الدافعية، الريادة، وبين التحصيل الأكاديمي لطلبة السنة الأولى؟

أما البحث المسحي، فإنه يستخدم أدوات مثل الاستبيانات والمقابلات لتجميع المعلومات من مجموعات من الأفراد. وهذه الأبحاث المسحية تتيح للباحثين تلخيص مميزات بمجموعات مختلفة لقياس اتجاهاتهم وآرائهم إزاء قضية ما. وتستخدم الأبحاث المسحية على نطاق واسع في التربية والعلوم الاجتماعية. فمثلاً، قد يسأل أحد الباحثين التربويين مجموعة من الآباء عن ماهية برنامج التربية الجنسية الذي يجب تقديمه لتلاميذ المدارس المتوسطة- إذا ما كان هنالك تعليم كهذا. إن مسحاً للمعلمين يمكنه الكشف عن اتجاهاتهم إزاء قبول أطفال معاقين في الفصول النظامية. واستطلاعات الرأي عبارة عن مسوحات تقام لتحديد من هم الناس المحتملين للتصويت لهم، أو المواقف التي يتخذونها إزاء قضايا معينة. وسوف نناقش البحوث المسحية في الفصل الثاني عشر.

البحث النوعي Qualitative Research

يسعى البحث النوعي لفهم أحد المواقف بالتركيز على الصورة الشاملة وليس على تفكيكه إلى متغيرات. والهدف هو صورة كلية شاملة وتعميق الفهم، وليس مجرد تحليل عادي للبيانات. فمثلا، لقد لوحظ لفترة طويلة بأن الاختلافات في الخلفية التربوية وحدها لا تبدو مسؤولة عن المصاعب التي يواجهها الطلبة السود في جامعة كانت مسبقاً للطلبة البيض. والسؤال هو: كيف يتصور الطلبة السود خبرتهم الأكاديمية في جامعة للطلبة البيض؟ وهذا السؤال يمكن سبر غوره

باستعمال منهجية البحث النوعي. فالباحث سوف يركز على عدد قليل من الطلبة السود، ثم يدرسهم بتفصيل كبير من خلال استخدامه الملاحظة والمقابلات المعمقة. وتتضمن نماذج معينة من البحث النوعي وصف الأجناس، ودراسة الحالة، وتحليلات المحتوى، وسوف نناقش البحث النوعي في الفصل الثالث عشر.

وصف الأجناس Ethnography: إن هذا العلم يشير إلى جمع البيانات حول أحد المواقف أو مجموعة ثقافية باستخدام المقابلات مكثفة الملاحظة. فمثلا، قد يود أحد الباحثين تقصي حجرة دراسية في مدرسة ابتدائية عادية تحوي بعض التلاميذ المعوقين. ويقوم الباحث بملاحظة تلاميذ الفصل لمدة طويلة من الزمن، فيجمع بيانات عن التلاميذ والمعلم من خلال مقابلات. ويتم الاحتفاظ بسجلات مطولة شاملة مكتوبة حول كل ما جرى داخل حجرة الدراسة.

دراسات الحالة Case Studies: تتضمن دراسة الحالة، نموذجياً، فحصاً شاملاً معمقاً لأحد الأفراد والذي فيه يسعى الباحث نحو وصف المجال الشامل لسلوك الفرد والعلاقة بين هذه السلوكيات مع تاريخ وبيئة ذلك الفرد. وقد قام العالم فرويد باستخدام مكثف لطريقة دراسة الحالة.

تحليلات المحتوى Content Analyses: وهناك نموذج آخر للبحث النوعي يشمل تحليلات المحتوى. ويركز البحث في هذا النموذج على تحليل المحتوى لوثيقة مكتوبة.

مراحل نموذجية في البحث Typical Stages in Research

هناك عدد من الأنشطة التي يشترك فيها كل الباحثين، بغض النظر عن الطريقة المعينة التي يختاروها في بحوثهم. ومع إننا سنركز على هذه النشاطات بتفصيلات أوسع في فصول لاحقة، إلا أن ما سيأتي هو خلاصة موجزة للخطوات المتخذة.

اختيار المشكلة Selecting a Problem

يبدأ الباحثون بسؤال يعتقدون أن له صلة بقضية لها نتيجة كافية لتسويغ البحث. ويجب أن تكون المشكلة سؤالاً لا يوجد له جواب جاهز، لكنه سؤال ينطوي على توفر وسائل لإيجاد الأجوبة له من خلال الملاحظة أو الاختبار العلمي. لنتأمل السؤال التالي:

هل ينال الأطفال الذين يتعلمون القراءة باستخدام الطريقة الكلية للغة، درجات أعلى، في اختبار تحصيل القراءة، من أطفال قد تعلموا القراءة باستخدام الطريقة البصرية(*). ويمكن بحث

(*) الطريقة البصرية (Sight-approach): طريقة في تعليم القراءة تطلب من المتعلم أن يقول ما يراه (المراجع).

هذا الموضوع تجريبياً بواسطة مقارنة الدرجات حسب معيار لتحصيل القراءة، لمجموعتين متكافئتين في كل شيء باستثناء أن مجموعة تعلمت القراءة باستخدام الطريقة الكلية والأخرى باستخدام الطريقة البصرية (Goodman, 1986).

وهنالك أسئلة متميزة يتعذر الإجابة عليها باستخدام الإجراءات العلمية. فمثلاً، هل التدريب باستخدام الطريقة الكلية للغة جيد للطلبة؟ هذا النوع من السؤال لا يمكن استقصاؤه علمياً دون معرفة معنى "جيد للطلبة" بالضبط، أو الكيفية التي تلاحظ بها أو تقاس بها "الجودة". فلا ينبغي أن يتضمن منطوق مشكلة البحث أية أحكام قيمية.

صياغة الفرضية Formulating a hypothesis

بعد تشخيص مشكلة البحث، تلي دراسة شاملة لجميع البحوث السابقة التي ربما قد جرت على مشكلة البحث. إن مراجعة هذه البحوث ذات الصلة ضرورية لتحقيق فهم أعمق للمشكلة، وإعداد خلفية ما لصياغة الفرضية أو الفرضيات الخاصة بالدراسة. (والفرضية هي بيان لتوقعات الباحث بخصوص العلاقات بين متغيرات الدراسة). وأحسن دليل على الإطلاق لفرضية ذكية هو تحليل دقيق للبيانات المتوفرة ذات الأثر على المشكلة. ولدى استخدام المثال أعلاه، يمكن للمرء بعد مراجعة دقيقة القيام بافتراض ما يلي: "تؤدي الطريقة الكلية للغة إلى تحصيل للقراءة أعلى مما تقدمه الطريقة البصرية". وتستخدم هذه الفرضية في بحث يستقصي العلاقات. وفي أنماط أخرى من البحث يمكن للشخص أن يبتدئ البحث بتوجيه سؤال ما. وهذا صحيح خصوصا في معظم البحوث المسحية. أما في البحث النوعي فالفرضيات تتولد عندما تتجمع البيانات خلال سياق الدراسة.

بالإضافة إلى صياغة فرضية ما، يقوم الباحثون كذلك في هذه المرحلة ببلورة تعاريف المصطلحات التي سيستخدمونها في الدراسة.

اختيار استراتيجية بحثية وتطوير الأدوات

Selecting Research Strategy and Developing Instruments

وباستخدام طريقة التفكير الاستنتاجي، تتحدد مضامين الفرضية المقترحة، وهذا يعني ما يجب ملاحظته إذا ما كانت الفرضية صحيحة. وإذا ما كان صحيحا أن تعلم القراءة باستخدام الطريقة الكلية للغة يؤدي إلى تحصيل للقراءة أعلى من تعليم القراءة بالطريقة البصرية، فآنذاك ينبغي على المرء أن يلاحظ أن درجات تحصيل القراءة للطلبة الذين يتعلمون القراءة بالطريقة الكلية للغة، أعلى مما حصل عليه أقرانهم المكافئون لهم والذين يتعلمون القراءة بالطريقة البصرية.

وبالتالي فإن هذه الوسيلة يتبعها تحديد العينة (الأفراد) التي سيجري تطبيق البحث عليها، وكذلك اختيار الطريقة الملائمة. وتحتاج مشكلات معينة طرقاً كمية، كما يمكن معالجة أخرى

بإحدى الاستراتيجيات النوعية. ويؤثر اختيار وسيلة البحث في تفاصيل تصميم الدراسة وفي طرائق الملاحظة أو قياسات المتغيرات. وقد تكون الأدوات المقننة لقياسات المتغيرات متوفرة مسبقاً، أو ربما يتحتم على الباحث أن يقوم بتطويرها.

تجميع وتفسير البيانات Collecting and Interpreting the Data

لابد من اختيار النتائج المستخلصة من فرضيات الدراسة. وبناء على ذلك، فإن هذه المرحلة تتضمن تجميع البيانات التي تشمل الأوجه الروتينية لاستخدام الأدوات، وحفظ السجلات وتحديد المواعيد، وهكذا. وعادة ما تأخذ هذه المرحلة، خلافاً للاعتقاد الشائع، وقتاً أقل بكثير من مراحل التخطيط السابقة للدراسة.

وبعد التجميع، يقوم المرء بتحليل البيانات العددية - إحصائياً عادة - لتحديد فيما إذا كان البحث قد أتاح دليلاً يدعم الفرضية. وفي الطريقة العلمية لا يقوم المرء بالادعاء ببرهنة الفرضية - لان ذلك معناه التعامل بمصطلحات الحقيقة المطلقة التي لا تُعدّ إحدى ميزات هذه الطريقة - وفقط يقوم المرء باستنتاج أن الدليل يدعم أو لا يدعم الفرضية.

والبحث النوعي عادة لا يقوم باختيار فرضيات سابقة. وعوضا عن ذلك، يقضي الباحثون النوعيون وقتا بالملاحظة، وبإجراء المقابلات، وبتجميع البيانات من أجل إحراز فهم أعمق للموقف الخاضع للدراسة. ولا توجد هنالك بيانات عددية لتحليلها، وبدلاً من ذلك تشمل البيانات أوصافا، وملاحظات، وانطباعات، وتسجيلات، وصور فوتوغرافية، وما إلى ذلك. وعندما تتجمع هذه البيانات يدرك الباحث ماهية الأسئلة المهمة.

تقرير النتائج Reporting the Results

يتوجب على الباحثين أن يجعلوا إجراءاتهم، واستنتاجاتهم، ونتائجهم متوفرة بشكل واضح يدركه كل من لهم اهتمام بالموضوع. وهذا الأمر يشمل عرضا موجزاً وواضحاً للخطوات المتبعة في الدراسة، بتفصيلات كافية حتى يتمكن أي شخص آخر من إعادة الدراسة.

وسوف تتم مناقشة كل من مراحل البحث السابقة بالتفصيل في فصول لاحقة. ومن المحتمل، أن يندر قيام الباحثين باتباع التسلسل بالضبط كما هو موصوف في الفقرات السابقة. وهذه الأنشطة غالبا ما تتشابك، وربما يكون هناك تحرك جيئة وذهاباً من مرحلة إلى أخرى.

أسئلة يطرحها الباحثون التربويون
Questions Asked By Educational Researchers

إن السؤال المحدد الذي اختير للبحث سيعتمد بالطبع على المجال الذي يهتم به الباحثون، وخلفيتهم، والمشكلة الخاصة التي تواجههم. وعلى أية حال، يمكننا تصنيف الأسئلة في البحوث

التربوية على أساس نظري، (وهذه تتعامل مع مبادئ أساسية)، أو على أساس عملي يجري تصميمها لحل مشكلات الموقف اليومي الآنية.

أسئلة نظرية Theoretical Questions

الأسئلة النظرية ذات الطبيعة النظرية تكون كالآتي: ما هذا؟ أي كيف يحدث هذا؟ أو لماذا يحدث هذا؟ وفي البحث التربوي فإن الأسئلة التي تبدأ بالكلمة "ما" تصاغ بتحديد أكبر، مثل: ما هو الذكاء؟ أو ما هو الإبداع؟ والأسئلة النموذجية التي تبدأ بالكلمة "كيف" هي: كيف يتعلم الطفل؟ أو، كيف تتطور الشخصية؟ أما الأسئلة التي تبدأ بالكلمة "لماذا" ربما تسأل كالآتي: لماذا ينسى المرء؟ أو، لماذا يكون أكثر توجهاً للإنجاز من الأطفال الآخرين؟.

قد يكون البحث ذو الاتجاه النظري موجهاً نحو تطوير النظريات أو نحو اختبار نظريات قائمة. ويشمل النوع الأول نمطاً في الدراسة يسعى الباحثون فيها لاكتشاف تعميمات حول السلوك بهدف توضيح طبيعة العلاقات القائمة بين المتغيرات. وقد يعتقدون بأن علاقة تقوم بين متغيرات معينة، وهكذا يجرون بحوثهم لوصف طبيعة تلك العلاقة. وفي ضوء ما يتوصلون إليه، قد يبدؤون بصياغة إحدى النظريات حول تلك الظاهرة. وقد تطورت نظريات التعلم بهذه الطريقة عندما أصبح الباحثون قادرين على إبراز العلاقات بين طرائق معينة، وبين الفرد والمتغيرات البيئية، وكفاية عملية التعلم.

وربما كانت الدراسات التي تهدف إلى اختبار النظريات القائمة فعلاً هي الأكثر شيوعاً في البحوث التربوية. ولربما يكون هناك طموح مفرط، وخاصة لدى الباحثين المبتدئين في التربية، باتخاذ هدف لهم وهو تطوير نظرية ما. والأمر الأكثر واقعية في العادة هو السعي نحو استنتاج فرضيات من نظريات التعلم القائمة، ومن الشخصية، ومن الدافعية وما شابه ذلك، ومن ثم القيام باختبار هذه الفرضيات. وإذا ما كانت تلك الفرضيات استنتاجات منطقية من النظرية، وقدمت الاختبارات التجريبية الدليل الذي يدعم الفرضيات، وكان ذلك الدليل داعماً للنظرية ذاتها.

أسئلة عملية Practical questions

إن العديد من الأسئلة في البحوث التربوية ذات طبيعة عملية مباشرة، تهدف إلى حل مشكلات معينة ربما تواجه تربويين خلال أنشطتهم اليومية. وهي أسئلة وثيقة الصلة بالبحث التربوي بسبب تعاملها مع مشكلات حقيقية في مستوى التطبيق. وهذه أمثلة على هذه الأسئلة: ما مدى تأثير اختيار بعض الطلبة لتعليم أقرانهم في أحد فصول المدرسة الابتدائية؟ أو، ما هو تأثير تعليم الأطفال استراتيجيات معرفية على إدراكهم لما يقرؤون؟ أو، ما هو التأثير لطريقة "مناقشة المشكلة" مقارنة مع طريقة "المحاضرة" في تدريس مادة الدراسات الاجتماعية لطلبة المرحلة الثانوية؟ إن الأجوبة على أسئلة كهذه قد تكون قيمة جداً في مساعدة المعلمين لاتخاذ قرارات عملية.

ويمكن بحث هذه الأسئلة العملية بطريقة علمية مثلما هو الحال مع المشكلات النظرية. إن النموذجين الاثنين من هذه الأسئلة يتميزان عن بعضهما على أساس الأهداف التي يودون تحقيقها وليس على صعيد التعقيد في الدراسة.

البحث الأساسي والبحث التطبيقي Basic and Applied Research

هناك نظام آخر للتصنيف يستخدم في بعض الأحيان مع البحث الذي يتعامل مع أسئلة عملية ونظرية، ويستند التصنيف على الهدف من البحث نفسه. إن النموذج الأول من البحث — والذي هدفه الحصول على البيانات التجريبية التي يمكن استخدامها في صياغة، وتوسيع، وتقييم النظرية — يدعى البحث الأساسي. وهذا النمط من الدراسة لا يتجه في التصميم أو الهدف نحو إيجاد حلول لمشكلات عملية. بل أن هدفه الأساسي هو توسيع آفاق المعرفة دون الالتفات إلى تطبيقات عملية. وبالطبع، قد تؤدي تلك الاستنتاجات إلى تطبيقها في نهاية الأمر في مشكلات عملية ذات قيمة اجتماعية. فالتقدم مثلا، في ممارسة الطب يعتمد على البحوث في ميدان الكيمياء الإحيائية وعلم الجراثيم. وبالمثل، فإن التقدم في الممارسات التربوية يعود إلى التقدم في اكتشاف قوانين عامة من خلال البحوث الأساسية في العلوم النفسية، والتربوية، والاجتماعية. والاهتمام الأولي للبحث الأساسي هو — على أية حال — اكتشاف المعرفة فقط من أجل المعرفة ذاتها. ولا تعيق غرضه هذا اعتبارات الفائدة الاجتماعية لتلك الاكتشافات.

والنموذج الثاني للبحث —الذي يهدف إلى حل مشكلة عملية آنية — يدعى البحث التطبيقي. ويتصل إجراء هذا البحث بمشكلات حقيقية وتحت ظروف وجودها خلال الممارسة. ومن خلال البحث التطبيقي، يكون بوسع التربويين حل مشكلاتهم على صعيد مناسب من التعقيد: ونعني بذلك الموقف التعليمي – التعلمي في حجرة الدراسة. وقد نعتمد على البحث الأساسي في اكتشاف القوانين الأكثر عمومية للتعلم، لكنه يجب على البحث التطبيقي أن يجرى من أجل تحديد الكيفية التي تعمل بها هذه القوانين داخل حجرة الدراسة. وهذه الطريقة مهمة إذا ما كانت التغيرات العلمية في الممارسات التعليمية ستتأثر. وما لم يأخذ المربون على عاتقهم حل مشكلاتهم الذاتية العملية، فلا يوجد هناك أحد يقوم بذلك. وتجب الإشارة هنا إلى أن البحث التطبيقي يستخدم أيضاً الطريقة العلمية في البحث.

ونجد أنه ليس هناك دائما حد فاصل واضح بين البحث الأساسي والبحث التطبيقي. وبالتأكيد فإن التطبيقات تنطلق من النظرية لكي تساعد في حل المشكلات العملية. ومن الجهة الأخرى، فلربما يعتمد البحث الأساسي على البحث التطبيقي لإتمام صياغاته النظرية. فأية تجربة تعلمية تجرى في حجرة الدراسة - على سبيل المثال - قد تساعد في اختبار نظرية تعلم معينة.

وفي الواقع، ففي السنوات الأخيرة كان هنالك تحرك باتجاه دمج البحثين الأساسي والتطبيقي. فمال البحث الأساسي نحو التوجه للدراسات الصفية. فعلى سبيل المثال، غالباً ما

يذهب الباحثون القائمون على تطوير نظريات عامة في التعلم إلى داخل الفصول الدراسية من أجل تفهم الكيفية التي يتعلم بها الأطفال، إذ يتوجب على الباحثين أن يولوا اهتماماً بمتغيرات مثل السياق والبنية الاجتماعية، ولا تتوفر هذه في البيئات الاصطناعية للمختبر.

لغة البحث العلمي THE LANGUAGE OF RESEARCH

يحتاج أي نظام علمي إلى لغة محددة لوصف وتلخيص الملاحظات في مجال ما. ويحتاج العلماء إلى مصطلحات على المستوى التجريبي لكي يصفوا ملاحظات معينة، كما يحتاج العلماء أيضاً إلى مصطلحات على المستوى النظري للإشارة إلى العمليات الافتراضية التي قد لا تخضع للملاحظة المباشرة. وربما يستخدم العلماء كلمات مأخوذة من اللغة اليومية، لكنهم غالباً ما ينسبون لها معاني جديدة خاصة لا توجد عادة في الاستخدامات العادية. أو ربما يتحتم عليها أحياناً إدخال مصطلحات فنية جديدة ليست جزءاً من اللغة اليومية. وغاية هذا المقطع، تعريف القارئ ببعض المصطلحات العامة المستخدمة في البحث التربوي.

والمصطلحات التي يستخدمها العلماء على الصعيدين الوصفي والنظري تشمل عناوين للمفاهيم والمفاهيم البنائية.

المفاهيم والمفاهيم البنائية Concepts and Constructs

"المفهوم" تجريد من أحداث نلاحظها، فهو كلمة تمثل التشابهات أو الأوجه العامة لموضوعات أو أمور تكون خلافاً لذلك، مختلفة بعضها عن بعض. فكلمات مثل "كرسي، كلب، شجرة، أو سائل" وآلاف أخرى من الكلمات في لغتنا تمثل مظاهر مشتركة لأشياء تختلف باختلاف تلك المظاهر. والهدف من مفهوم ما هو تبسيط التفكير بواسطة شمول عدد من الأمور تحت عنوان عام واحد. وتكون بعض المفاهيم قريبة جداً من الأمور التي تمثلها. وهكذا – مثلاً – فإن معنى المفهوم "شجرة" يمكن أن توضحه بسهولة وذلك بالإشارة إلى أشجار محددة. فالمفهوم تجريد للميزات التي تشترك فيها كل الأشجار — مميزات يمكن ملاحظتها مباشرة.

وعلى أية حال، فإن مصطلحات مثل "الدافعية، والعدالة، وقدرة حل المشكلات" لا يمكن توضيحها بسهولة بالإشارة إلى حدث أو شيء ما. وهذه التجريدات ذات المستوى العالي تدعى "مفاهيم بنائية". فلقد جمع الناس أو قاموا ببناء تجريدات أكثر تعقيداً من مفاهيمها. وبنفس الطريقة التي نشيد بها داراً ما وذلك بواسطة وضع الخشب والمواد الأخرى سوية في نمط ذي غاية مقصودة، فإننا نخلق مفاهيم بنائية بواسطة دمج مفاهيم ومفاهيم بنائية أقل تعقيداً في أنماط هادفة. فمثلاً، مفاهيم مثل حدة النظر، والتمييز بين الرموز، والتوجه من اليسار إلى اليمين، ومفردات الإصغاء، وأخري غيرها يجري دمجها بطريقة هادفة لإنتاج المفهوم البنائي "الاستعداد

للقراءة". والمفاهيم البنائية مفيدة في تفسير البيانات التجريبية وفي بناء النظريات. وتستخدم المفاهيم البنائية للتعبير عن الانتظامات الملحوظة وعن العلاقات.

ويتم إنشاء المفاهيم البنائية من أجل تلخيص الملاحظات ولتقديم التفسيرات. ويتم التخلي عن المفهوم البنائي عندما توجد طريقة أفضل، لتفسير وتلخيص الملاحظات تحل مكانه. فمثلاً، الملاحظات بأن 1- بعض المواد تحترق وأخرى لا تحترق، 2- بعض المواد تحترق بشدة أكثر من مواد أخرى، جرى تلخيصها بالمفهوم البنائي "فلوجستن: وهي مادة كيميائية وهمية كان يعتقد بأنها قبل اكتشاف الأوكسجين، مقوم أساسي من مقومات الأجسام الملتهبة". وقد جرى الاستغناء عن هذا المفهوم البنائي عندما تطورت تفسيرات أكثر فائدة لعملية الاحتراق.

تحديد المعنى Specification of Meaning

كلما ابتعدت المفاهيم أو المفاهيم البنائية لشخص ما عن الحقائق التجريبية أو عن الظواهر المقصود أن تمثلها، ازدادت إمكانية سوء الفهم وكذلك الحاجة إلى تعريفات دقيقة. والمعاني الدقيقة للكلمات الموجودة في مفردات العالم يتحتم ترسيخها. فيتوجب تعريف المفاهيم بمصطلحات مجردة تنقل المعنى العام الذي يفترض أنها تحويه، وبمصطلحات الإجراءات التي ستقاس بها أو ستفعّل بها في دراسة معينة. والنموذج الأولي للتعريف يدعى تعريفاً تأسيسياً، والنموذج اللاحق يسمى بالتعريف الإجرائي.

التعريف التأسيسي Constitutive Definition

التعريف التأسيسي هو تعريف رسمي يجري فيه تعريف مصطلح باستخدام مصطلحات أخرى. فمثلاً، يمكن تعريف الذكاء على أنه القدرة على التفكير تجريدياً. إن هذا النمط من التعريفات يساعد في إيصال الطبيعة العامة للظاهرة التي يهتم بها الباحث، كما أنه أيضاً يبين علاقتها، بدراسات أخرى تستخدم مفاهيم مشابهة وبالنظرية. والتعريف التأسيسي يشرح المصطلح وربما يقدم للمرء بعض الفهم للظاهرة التي يصفها المصطلح. وعلى أية حال، إذا كان على شخص ما أن يجري أحد البحوث، فيتحتم على ذلك الشخص ترجمة المفاهيم إلى أحداث قابلة للملاحظة.

التعريف الإجرائي Operational Definition

إن "التعريف الإجرائي" هو تعريف ينسب معنى إلى مفهوم أو مفهوم بنائي وذلك بتحديد الإجراءات التي يتوجب القيام بها من أجل قياس أو تفعيل المفهوم. وهذا النوع من التعريف حيوي في البحث العلمي لأنه ينبغي جمع البيانات بلغة أحداث قابلة للملاحظة. ويمكن للعلماء التعامل مع مفاهيم بنائية مثل: التعلم، الدافعية، القلق، أو الإنجاز، على صعيد نظري، ولكن قبل أن يقوم العلماء بدراستها تجريبياً، يتوجب عليهم تحديد أحداث قابلة للملاحظة تمثل تلك المفاهيم البنائية. ويختار العلماء أحداثاً مميزة كمؤشرات للمفاهيم المجردة ثم يصممون إجراءات

للحصول على بيانات ذات صلة بالمفهوم، وذلك لتعريف مفهوم أو مفهوم بنائي إجرائياً.

وهناك نموذجان للتعريفات الإجرائية، مُقاس، وتجريبي. والتعريف الإجرائي المقاس يشير إلى الإجراءات التي يقوم بها الباحثون لقياس مفهوم. فمثلا، ربما يعرف الذكاء إجرائياً على أنه الدرجات الناتجة عن "اختبار ستانفورد – بينيه للذكاء"، أو يمكن للإبداع أن يشير إجرائياً إلى الدرجات الناتجة عن "اختبار مينيسوتا للإبداع". ويشير التعريف الإجرائي إلى الخطوات التي يتخذها أحد الباحثين لإحداث ظروف تجريبية معينة. فمثلا، التعريف الإجرائي للإحباط في دراسة بحثية يمكن أن يأخذ شكلا يعيق الفرد من الوصول إلى أحد الأهداف، أو ربما يشير الإدراك المتنافر، إجرائيا إلى إلزام أفراد عينة الدراسة بالتصريح ببيانات علنية مخالفة لاعتقاداتهم الخاصة.

وبالرغم من أن الباحثين يهتدون بمعرفتهم وخبرتهم وتقارير باحثين آخرين، إلا أن التعريف الإجرائي يظل – إلى حد ما- إجراءً كيفياً. وغالبا ما نختار من مجموعة منوعة من التعريفات الإجرائية الممكنة تلك التي تمثل بأحسن صورها طريقتنا في التعامل مع المشكلة. وبالتأكيد، فإن التعريف الإجرائي لا يستنفذ المعنى العلمي التام لأي مفهوم. إنه محدد جداً في المعنى، وغايته هي وضع حدود للمصطلح، وذلك من أجل التأكيد بأن أي فرد يهمه الأمر يفهم الطريقة الخاصة التي يستخدم بها المصطلح. وتعتبر التعريفات الإجرائية كافية إذا ما كانت إجراءاتها تجمع بينات تشكل مؤشرات مقبولة للمفاهيم المجردة المقصود تمثيلها. وغالبا ما تكون القضية، وجهة نظر، تحدد فيما إذا تم إنجاز النتيجة أم لا.

إن التعريفات الإجرائية أساسية للبحث لأنها تتيح للباحثين قياس المفاهيم المجردة والمفاهيم البنائية، كما تسمح للعلماء بالتحرك من مستوى المفاهيم البنائية والنظرية إلى مستوى الملاحظة التي يرتكز عليها العلم. ويتمكن الباحثون من الاستمرار في استقصائهم التي يتعذّر قيامها دون استخدام التعريفات الإجرائية. ومن المهم أن يتذكر المرء انه، بالرغم من قيام الباحثين بنشر استنتاجاتهم بلغة المفاهيم البنائية المجردة وربطها ببحوث ونظريات أخرى، فإن ما وجده الباحثون فعلاً هو علاقة، بين مجموعتين من بيانات تقبل الملاحظة والقياس، تم اختيارها لتمثيل المفاهيم البنائية. فالبحث في العلاقة بين المفهومين البنائيين (الإبداع والذكاء) سوف ينسب بصورة عملية، الدرجات في اختبار للذكاء إلى الدرجات في مقياس للإبداع.

المتغيرات Variables

يقوم الباحثون بدراسة المتغيرات والعلاقات القائمة بينها. والمتغير يدل على، أو يعبر عن، مفهوم أو مفهوم بنائي. ويكتسب المتغير قيماً مختلفة. فالطول هو مثل واحد على متغير ما، ويمكنه أن يتغير في فرد ما من وقت لآخر، وبين الأفراد في وقت واحد، وبين المتوسطات في المجموعات، وهكذا. فالطبقة الاجتماعية، والجنس، ومستوى المفردات، ومعامل الذكاء، ودرجات اختبار

التهجي، كلها جميعا أمثلة أخرى على المتغيرات. وفي دراسة اهتمت بالعلاقة الموجودة بين مستوى المفردات اللغوية وبين التحصيل لتلاميذ السنة الثامنة، فإن المتغيرات ذات الأهمية سوف تكون قياسات المفردات اللغوية وقياسات التحصيل العلمي. وهناك طرائق مختلفة لقياس التحصيل العلمي. وبوسع الباحث استخدام اختبار تحصيل مقنن، أو اختبار يعده المعلم، أو درجات في صف علمي، أو تقييمات لمشاريع علمية منجزة. وكل مقياس من هذه المذكورة آنفا يمكن استخدامه لتمثيل المتغير "التحصيل العلمي". وعكس المتغير هو "الثابت". والثابت هو قيمة لا تتغير داخل الدراسة. ومفهوم مستوى السنة الدراسية — بالرغم من كونه متغيراً بالتعريف، هو في هذه الدراسة أحد "الثوابت" بسبب أن جميع الأفراد هم تلاميذ السنة الثامنة. وفي دراسة تقارن حالات طالبات في الدراسة الثانوية يخططن لمستقبل مهني وبين أخريات لا تخططن، فإن "طالبات الدراسة الثانوية" هو "الثابت" أما المتغير المستقل فهو التخطيط (أو عدمه) لمستقبل مهني، أما الاتجاهات فتمثل المتغير التابع.

أنواع المتغيرات Type Of Variables

هناك عدة طرق لتصنيف المتغيرات، فالمتغيرات يمكن أن تكون تصنيفية وتسمى نوعية/ طبقية أو تكون مستمرة. وعندما تكون العينات البشرية مصنفة وذلك بتقسيمها إلى مجموعات، فإن الصفة التي يستند إليها التصنيف يصطلح على تسميتها " المتغير النوعي". إن لغة البيت، والمحافظة التي يتم فيها السكن، والعمل الأساسي للوالد، والمدرسة التي درس فيها، جميعها أمثلة على متغيرات نوعية. إن ابسط نموذج للمتغير النوعي له صنفان فقط، ويدعى "المتغير الثنائي". فالذكر/ والأنثى، والمواطن/ والأجنبي، والنجاح/ والرسوب، جميعها متغيرات ثنائية. ولبعض المتغيرات النوعية أكثر من صنفين، وبعض الأمثلة هي المستوى التعليمي، والانتماء الديني، وبلد المولد.

وعندما تمتلك صفة مميزة عدداً لا منته من القيم ضمن المجال، عندئذ تسمى "متغيراً مستمراً/ متصلاً". فعندما يزداد طول طفل من 100 سم إلى 105 سم فإن طوله يمر بعدد لا منته من الأطوال. فالطول والوزن والعمر ودرجات التحصيل أمثلة على متغيرات مستمرة. إن أهم تصنيف للمتغيرات على الإطلاق يستند على "استخدامها" في البحث، وذلك عندما يجري تصنيفها كمتغيرات مستقلة أو متغيرات تابعة.

وهناك بعض المتغيرات السابقة لمتغيرات أخرى. ويمكننا معرفة ذلك تجريبياً، أو ربما نفترض بالاستناد إلى إحدى النظريات بأن أحد المتغيرات سابق لآخر. وعلى سبيل المثال، من أجل أن يتمكن شخص ما من القراءة، فإنه يفترض بأن ذلك الشخص يحتاج بعضاً من الذكاء، وهذا معناه: أن متغير "الذكاء" هو سابق لمتغيّر "القراءة". وإلى درجة ما، فإن القدرة القرائية هي نتيجة لمتغير "الذكاء" فهي تعتمد على مستوى ذكاء الفرد. وخلال البحث، فإن المتغيرات التي هي نتيجة أو معتمدة على متغيرات سابقة لها تدعى "متغيرات تابعة". والمتغيرات السابقة لمتغير تابع تسمى "متغيرات مستقلة". فمثلا، إن طول الطفل (متغير تابع) يعتمد إلى درجة ما على عمره (متغير مستقل).

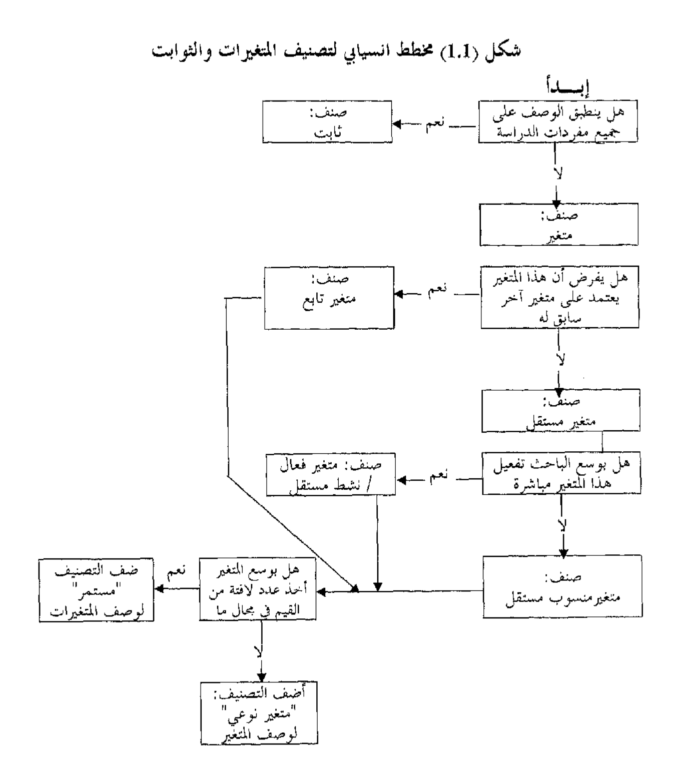
إن المتغير التابع هو الظاهرة وهو هدف الدراسة والبحث، أما المتغير المستقل فهو عامل منفصل متميز عن المتغير التابع بصورة ملحوظة، لكنه ربما تكون له صلة بالمتغير التابع. وقد تعمل عوامل كثيرة كمتغيرات مستقلة – مثل الطبقة الاجتماعية، والبيئة البيتية، وظروف غرفة الصف – هي أوجه متميزة للبيئة. بالإضافة إلى ذلك، فإن صفات مميّزة شخصية مثل العمر، الجنس، الذكاء، والدافعية، قد تكون متغيرات مستقلة لها صلة بالمتغير التابع.

ولاحقاً، عندما نناقش الطرق التجريبية للبحث، سنعرف المتغير المستقل كمتغير يقوم الباحثون التجريبيون بتفعيله (التأثير عليه) أو تغييره. والمتغير الذي تظهر عليه آثار التفعيل (المعالجة) بالملاحظة يدعى المتغير التابع. وهو يسمى كذلك لأن قيمته تعتمد – أو من المفترض اعتمادها – على قيمة المتغير المستقل. فمثلا، من أجل دراسة تمرين مستند إلى الحاسوب في تحصيل الرياضيات، يعالج الباحثون طريقة التعليم (متغير مستقل) وبعد ذلك يلاحظون التأثير على تحصيل الرياضيات (متغير تابع). وبعد إقامة العلاقة بين المتغيرات من خلال البحث، قد يتنبأ شخص انطلاقاً "من" متغير مستقل "إلى" المتغير التابع. وفي البحوث التربوية، فإن طرق التدريس والأساليب من المحتمل لها أن تكون أكثر استخداماً لمتغيرات مستقلة. والأخرى تشمل: العمر، الجنس، الطبقة الاجتماعية، الاتجاهات، الذكاء، والدافعية والإبداع. والتحصيل المدرسي أو التعلم هما الأكثر شيوعاً كمتغيرات تابعة.

ومن الممكن لأحد المتغيرات أن يكون متغيرا مستقلاً في دراسة ما، ومتغيراً تابعا في دراسة أخرى. وفيما إذا كان المتغير يعتبر مستقلاً أو تابعا فإن ذلك تحدده الغاية من الدراسة. وإذا ما جرى بحث تأثير الدافعية على التحصيل، عند ذلك نعتبر الدافعية متغيراً مستقلاً. وبالرغم من هذا كله، إذا ما رغب شخص ما في دراسة أثر أساليب الاختبارات، أو ترتيبات تجميع طلبة الفصل في مجموعات، أو أساليب وضع الدرجات على دافعية الطلبة، عندئذ يصبح الدافعية المتغير التابع. والذكاء عموما يعتبر متغيرا مستقلاً بسبب اهتمامنا بتأثيره على التعلم الذي يعتبر متغيراً تابعاً. وعلى أية حال، ففي دراسات تقوم ببحث تأثير خبرة السنوات الدراسية للروضة على التطور العقلي للأطفال، يكون الذكاء متغيرا تابعا. ويمكن القيام بتصنيف المتغيرات المستقلة على أساس فيما إذا كان بوسع الباحثين تفعيل المتغير المستقل. وفي هذا التصنيف، هناك نموذجان اثنان من المتغيرات المستقلة: فعّال/ نشيط ومنسوب. ويعّرف المتغير الفعال بأنه متغيّر يمكن للباحثين تفعليه على نحو مباشر. فمثلا، طريقة التعليم، طريقة التقسيم إلى مجموعات، وتعزيز الأساليب، كلها متغيرات يمكن تفعيلها، ولهذا تدعى متغيرات فعّالة. أما المتغير المنسوب فلا يتمكن الباحثون من تفعيله بصورة نشيطة. إن متغيرات كهذه – وتدعى بعض الأحيان "متغيرات مفروضة" هي مميزات لأفراد لا يمكن تفعيلها متى نشاء. فمثلا اللياقة، الجنس، العرق، العمر، والطبقة الاجتماعية هي متغيرات منسوبة نموذجية. ويتمكن الباحثون من دمج متغيرات منسوبة ببحوثهم وذلك بتحديد الأفراد داخل مجموعات على أساس متغيرات موجودة قبليا.

وفي بعض الحالات ليس في الإمكان معرفة المتغير الذي يؤثر على متغير ثانٍ. فمثلا عندما يجري بحث العلاقة بين الجزمية (الدوغماتية) والاتجاه السياسي، لا يمكن للفرد تحديد أي من المتغيرات يكون مستقلاً أو تابعاً. وهذا النوع من المواقف لا يقدم يد العون للتجريب. ويجري البحث عادة في العلاقة بين متغيرات كهذه في دراسات الارتباط. والشكل رقم (1.1) يبين إحدى الطرائق لتصنيف المتغيرات والثوابت.

شكل (1.1) مخطط انسيابي لتصنيف المتغيرات والثوابت



الخلاصة SUMMARY

يحاول الإنسان الإجابة على أسئلة ما من خلال الخبرة، والمرجعية، والتفكير الاستنتاجي، والتفكير الاستقرائي، والطريقة العلمية. وكل وسيلة تحتاج إلى افتراضات معينة. وصحة الأجوبة تعتمد على صحة الافتراضات القائمة كأساس للطريقة المتبعة.

ترتكز الطريقة العلمية على افتراضين أساسيين: 1- يمكن اشتقاق الحقيقة من الملاحظة، 2- تخضع الظواهر لعلاقات طيّعة للقانون (*).

ولا ينشد الباحثون الحقيقة المطلقة، بل ينشدون نظريات تسعى إلى تفسير الظواهر والتنبؤ بوقوعها بطريقة موثوقة. إنها تبحث عن نظريات تكون اقتصادية، وقابلة للاختبار، وتتميز بالثبات، وكذلك عن نظريات هي في حد ذاتها مثيرات لبحوث لاحقة. والطريقة العلمية تلتزم بتصحيح ذاتها، كما أن كل نظرية تعتبر مؤقتة/ تجريبية ، وربما يجري التخلي عنها إذا ما برزت نظرية جديدة تلائم المعايير على نحو أفضل.

لقد استخدمت الطريقة العلمية لقرون عديدة لتفسير، وتوقع، والسيطرة على، الظواهر الطبيعية ، لكنها استخدمت نسبيا في التربية مؤخراً. إن تعقيد المتغيرات التربوية والصعوبات التي تواجهها الملاحظات الجديرة بالثقة أعاقت البحث العلمي في التربية. وعلى أية حال، فمنذ بداية الحركة في مطلع القرن الماضي، تمتع البحث العلمي في التربية بقبول متزايد ونجاح متعاظم في كل من البحوث النظرية والعلمية.

المنهجان الأساسيين في البحث هما الكمي والنوعي. فالبحث الكمي الذي يعالج أسئلة مثل ما مدى حجم وما جودة، يمكن علاوة على ذلك، تصنيفه كتجريبي وغير تجريبي. والبحث النوعي يهتم بالحصول على فهم تام لما يحدث في وضع معين. أما البحث التاريخي الذي يركز على دراسة أحد الأحداث الماضية فهو نموذج آخر من منهج البحث.

أما الخطوات النموذجية في أي بحث تربوي فهي: 1- اختيار المشكلة، 2- صياغة الفرضية، 3- انتقاء استراتيجية البحث وتطوير الأدوات، 4- تجميع وتحليل وتفسير البيانات، 5- تبادل الاستنتاجات بنشر نتائج الدراسة.

واستنادا إلى الهدف، فإن للبحث التربوي صنفين رئيسين: أساسي وتطبيقي. والاهتمام الأولي للبحث الأساسي هو تمديد وتوسيع حدود المعرفة واكتشاف قوانين عامة. أما الهدف الرئيسي للبحث التطبيقي فهو حل المشكلات العملية الآنية.

(*) يشار إلى هذا الافتراض أيضا "بالحتمية العامة" التي تقول بأن لكل ظاهرة أسبابها، فهو يؤكد أن الظواهر تحكمها قوانين تنظم علاقات بين العلل والمعلولات (الظواهر) "المراجع".

وعلى مستوى نظري، فإن العلماء التربويين يستخدمون مصطلحات مثل الذكاء والإبداع والقدرة على حل المشكلة والدافعية، التي هي تعبيرات تجريدية جاءت من ملاحظة سلوك معين. وهذه تدعى مفاهيم بنائية. وفي البحث الكمي، فإنه يجري "تكميم"(*) المفاهيم البنائية لتأخذ قيماً مختلفة. وهكذا تدعى "متغيرات". وهناك نموذجان رئيسيان للمتغيرات: المستقل، والتابع. وإذا ما كان متغيّر سابقاً لمتغيّر آخر، عندئذ يدعى "المتغير المستقل"، لكنه إذا ما كان نتيجة لمتغير آخر، يسمى حينئذ "المتغير التابع".

وفي البحث، فإن تعريف المتغيرات بصورة إجرائية أمر أساسي. وهنالك طريقتان لتعريف المتغيرات بصورة إجرائية: 1- باستعمال نوع ما من المقاييس، 2- أو بواسطة تحديد الخطوات المتخذة في تجربة ما لتوفير ظروف بحث معينة. والنوع الأول يدعى "تعريفاً إجرائياً مقاساً" والثاني يسمى "تعريفاً إجرائياً تجريبياً".

مفاهيم أساسية Key Concepts

متغير مستقل فعال/ نشط	active independent variable
بحث تطبيقي	applied research
متغير مستقل منسوب	attribute independent variable
بحث أساسي	basic research
متغيرات نوعية	categorical variables
مفهوم	concept
ثابت أو لا يتغير	constant
تعريف تأسيسي	constitutive definition
مفهوم بنائي	construct
متغير مستقل/ متصل	continuous variable
الاستنتاج	deduction
متغير تابع	dependent variable
متغير ثنائي	dichotomous variable
تعريف إجرائي تجريبي	experimental operational definition
فرضية	hypothesis
استقراء ناقص	imperfect induction
متغير مستقل	independent variable
استقراء	induction

(*) تكميم (Quantification): مصطلح يشير إلى قياس الكمية، أي إجراء قياس كمي لصفة ما. (المراجع)

measured operational definition	تعريف إجرائي مقاس
operational definition	تعريف إجرائي
parsimony principle	مبدأ الاقتصاد/ التقتير
scientific approach	الطريقة العلمية
theory	نظرية
variable	متغير

تمارين EXERCISES

1- عين مصدر المعرفة – تفكير استنتاجي، تفكير استقرائي، أو الطريقة العلمية – المستخدمة بصورة واضحة جدا في الأمثلة التالية:

أ – بعد ملاحظة طويلة ومكثفة للتفاعلات، استنتج لافوازيه بأن الاحتراق عملية تقوم فيها مادة محترقة بالاتحاد مع الأوكسجين. وكان عمله صفعة الموت لنظرية الاحتراق القديمة المسماة "فلوجستن".

ب- استنتج دالتون – بعد تفكير طويل – بأنه لاشك أن المادة تتكون من جزئيات صغيرة تدعى الذرات. وكانت افتراضاته الأولى المبكرة هي الأساس للنظرية الذرية.

ج- بعدئذ، تبنى العلماء افتراضات دالتون، وتوصلوا إلى استنتاجات منها، ثم انطلقوا في جمع بيانات أكدت تلك الافتراضات. وبهذا تم إيجاد الدعم للنظرية الذرية.

د- بإدراك أن المواد المشعة تطلق باستمرار جزيئات من الطاقة دون التقليل من كتلتها على ما يبدو، قام آينشتاين بتطوير صيغة يحول بواسطتها المادة إلى طاقة: $(E=MC^2)$.

هـ- بقبوله نظرية آينشتاين، قام العالم (فيرمي) بتجربة نتج عنها تفتيت الذرة.

و- بعد دراسته لنظرية التعزيز، تفترض إحدى المعلمات بأن استخدام برنامج تعليمي للحاسوب سوف يؤدي إلى تحصيل رفيع في الحساب. بعد ذلك تُصمم تلك المعلمة دراسة تستخدم فيها البرنامج التعليمي المذكور مع فصلين من تلامذة الصف السادس، بينما يجري استعمال مواد علمية تقليدية مع فصلين آخرين من تلامذة الصف السادس.

2- ما هو دور النظرية في الاستقصاء العلمي؟.

3- ما هو الفرق بين النظرية الاستقرائية والنظرية الاستنتاجية؟

4- صنف البحث التالي أما "أساسيا" أو "تجريبيا" بالاعتماد على عنوان كل دراسة:

أ – تأثير حقنات RNA (Ribonucleic Acid) على نقل المهارات من حيوانات مدربة

إلى حيوانات ليست مدربة.

ب- نتائج برنامج علاجي في مادة الحساب.

ج- الاشتراط / التطويع / التكييف كدالة للفترة بين المثيرين الشرطي (التطويعي) والأصلي (غير الشرطي أو الطبيعي).

د - تدريس الهندسة لرعاية التفكير التأملي: دراسة تجريبية.

5- صنف المتغيرات التالية، باستخدام (فعال، منسوب، أو كليهما) :

أ - كمية من الدواء المعطى.

ب- القلق.

ج- خلفية اقتصادية / واجتماعية.

د - طريقة تعليم.

6- ما هي الاتجاهات المتميزة للعلماء التي جرى انتهاكها في العبارات التالية:

أ - جرى الاضطلاع بهذه الدراسة للبرهنة على أن استعمال الماريجوانا مؤذ وضار بالإنجاز الأكاديمي.

ب- لقد برهنت بشكل نهائي بأن هذه هي الحالة.

ج- تظهر النتائج بأن الماريجوانا شريرة.

7- ما هي مميزات النظرية المفيدة؟

8- حدد ما يلي: في دراسة مصممة لتحدد تأثير المقادير المتغيرة للحرمان من النوم على تعلم المقاطع التافهة للكلمات:

أ - ما هو المتغير المستقل؟

ب- ما هو المتغير التابع؟

9 - صنف المتغيرات الآتية: أي منها نوعي، وأي منها مستمر؟

أ - التحصيل.

ب- الطريقة الصوتية في القراءة(*) مقابل طريقة أنظر/ واقرأ (الطريقة البصرية) في القراءة.

ج- تكلم الإسبانية، تكلم الإنكليزية، تكلم الفرنسية.

د - قوة العضلات.

هـ- الاستعداد الموسيقي.

10- ما هي مميزات التعاريف الإجرائية؟

(*) الطريقة الصوتية في القراءة: طريقة تنتقل من الحرف إلى الكلمة وتعتمد على صوت نطق الحرف (المراجع).

11- أي نوع من التعاريف الإجرائية(مقاس أو تجريبي) أكثر ملاءمة لكل من المتغيرات الآتية:
أ - التعزيز.
ب- التحصيل.
ج- الاتجاه.
د - طريقة التدريس.

12- أنت معلم كنت مهتما بمقدار السلوك العدواني الذي لاحظته بين الأطفال. ولقد قابلت الآباء أو الأمهات حول ممارساتهم في تربيتهم لطفلهم (استعمال العقاب، الثواب، وما شابهها)، ومقدار ما يشاهده الأطفال من التلفاز. ولقد جمعت بيانات كثيرة، لكنك تشعر بأن ليس هناك موضوع يوحدها. فما هي خطوتك المقبلة الأكبر احتمالا حسب الطريقة العلمية؟

13- أي وسيلة بحث (تجريبية، عليّة-مقارنة مسحية) ستعطيك أجوبة ذات تأثير كبير على كل سؤال؟
أ- هل الأطفال الذين يتناولون فطورهم يحصلون على درجات أفضل في المدرسة؟
ب- هل تقوم وحدة في الغذاء المناسب بتبديل عادات الأطفال في تناول فطورهم؟
ج- ما عدد الأطفال في المدرسة الذين يفصحون الإدارة بعدم تناولهم فطورا في البيت؟
د - هل برنامج مؤسسة الفطور مجاني في المدرسة له تأثير مختلف في أداء الطلبة؟

14- صنف كلا من الدراسات التالية طبقا للمنهج البحثي الأعلى احتمالاً في الاستخدام بالاعتماد على العناوين:
أ - الأداء الفارقي بين الجنسين لمفردات في الرياضيات.
ب- تأثيرات استخدام المعلومات الراجعة في الحاسوب على جدية الأخطاء اللاحقة.
ج- وجهات نظر وتقديرات طلبة الكلية للأستاذ المثالي.
د - تأثير غياب الأب في وقت مبكر على الاستعداد المدرسي.
هـ- مدرسة ثانوية بديلة: دراسة معمقة.

15- قدم أمثلة على استخدام الخبرة والمرجعية كمصادر للمعرفة.

16- أعط مثالا على كيفية قيام البحوث الأساسية في العلوم الأحيائية بتحسين الممارسات الطبية.

17- أعط مثالاً على كيفية قيام البحث الأساسي في التعلم بتحسين ممارسة التدريس.

18- أعط مثالاً عن بحث تطبيقي قد تم في حقل اهتمامك. اكتب قائمة بأسماء بحوث إضافية يحتاجها حقلك. ما المتغيرات التي يمكن استقصاءها في هذه الدراسات؟

19- ما هو أنسب منهج بحثي في اعتقادك يكون مناسباً لبحث كل سؤال بحثي من الأسئلة التالية:
أ - هل التعلّم التعاوني يعزز التعلم في فصول الكليات؟
ب- ما هي العلاقة بين التحصيل في القراءة والمفردات في سنوات الدراسة الابتدائية؟
ج- كيف يشعر الآباء والأمهات حول إدخال مادة "نقص المناعة المكتسب" في منهج الدراسة المتوسطة؟
د - ما هي العلاقة بين خلفية المعلمين أثناء دراستهم في الكلية (التربية مقابل الفنون العقلية)(*) وبين مقاييس معينة للكفاية داخل الصفوف؟
هـ- كيف آلت الإصلاحات التربوية في المرحلة المتوسطة إلى تطوير المدارس المتوسطة؟
و- ما هو تأثير التكوين العائلي على تكييف الأطفال مع دور الطالب؟

الأجوبة ANSWERS

1- أ. تفكير استقرائي.
ب. تفكير استنتاجي.
ج. الطريقة العلمية.
د. تفكير استنتاجي.
هـ. الطريقة العلمية.
و. الطريقة العلمية.

2- النظرية توحد الاستنتاجات، وتلخص المعلومات، وتقدم إرشادات لبحوث جديدة، كما تمكننا من تفسير وتوقع الظواهر.

3- النظرية الاستقرائية تساعد في تفسير ملاحظات سابقة، بينما النظرية الاستنتاجية قد تم بناؤها قبل إجراء الملاحظات الطويلة المكثفة.

4- أ. أساسي.
ب. تطبيقي.
ج. أساسي.
د. تطبيقي

(*) تشير التربية هنا إلى العلم الذي يبحث في أصول تنمية الفرد والمناهج والمبادئ والأهداف لهذه التنمية، أما الفنون العقلية فتشير إلى الدراسات النظرية كالعلوم واللغات وكان محورها الفلسفة التقليدية/ المثالية (المراجع).

5- أ. نشيط/ فعال
ب. يمكن أن يكون أي واحد منهما.
ج. منسوب.
د. نشيط.

6- أ. العالم يكون موضوعيا وغير متحيز.
ب. العالم شكاك، ويعتبر الاستنتاجات تجريبية/ مؤقتة.
ج. العالم يتعامل مع الحقائق، وليس مع القيم.

7- النظرية المفيدة تشرح الظواهر بأبسط الأشكال الممكنة. كما أنها تنسجم مع الملاحظات ومع المعرفة السائدة في حينه، وتقدم الوسائل من أجل التحقق منها، كما أنها تحفز على قيام بحوث جديدة.

8- أ. مقدار الحرمان من النوم.
ب. عدد المقاطع التافهة التي جرى تعلمها.

9- أ. مستمر/ متصل.
ب. نوعي.
ج. نوعي.
د. مستمر.
هـ. مستمر.

10- التعاريف المقبولة تحدد إجراءاً واضح المعالم لتكوين أو تحديد وجود الظواهر ومداها.

11- أ. تجريبي
ب. مُقاس
ج. مُقاس
د. تجريبي

12- حاول صياغة إحدى نظريات العدوان لدى الأطفال.

13- أ. علّي - مقارن.
ب. تجريبي
ج. مسحي
د. تجريبي

14- أ. علّي - مقارن
ب. تجريبي

ج. مسحي
د. علّي – مقارن
هـ. نوعي

15- الإجابات متنوعة.

16- الإجابات متنوعة.

17- الإجابات متنوعة.

18- الإجابات متنوعة.

19- أ. تجريبي
ب. الارتباطي
ج. مسحي
د. علّي – مقارن.
هـ. تاريخي
و. علّي – مقارن

REFERENCES المصادر

Darwin, F. (Ed.). (1899). *The life and letters of Charles Darwin* (Vol. 1). New York: Appleton.

Goodman, K. (1986). *What's whole in whole language.?* Portsmouth, NH: Heinemann.

Kerlinger, FN. (1986). *Foundations of behavioral research.* New York: Holt, Rinehart and Winston.

Mees, C.E.K. (1934). Scientific thought and social reconstruction. *General Electric Review,* 37, 113-119.

Pirsig, R.M. (1974). *Zen and the art of motorcycle maintenance: An inquiry into values. New* York: Morrow.

Rice, J.M. (1912). *Scientific management in education.* New York: Hinds, Noble and Eldredge.

Terman, L.M. (1926). The mental and physical traits of a thousand gifted children. *In Genetic studies of genius* (Vol. 1). Stanford, CA: Stanford University Press.



خلفية البحث

RESEARCH BACKGROUND

* مشكلة البحث.
* مراجعة الأدبيات.
* الفرضية.



مشكلة البحث

The Research Problem

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون بوسع الطالب:

1- تعريف مشكلة البحث

2- تشخيص مصادر محتملة لمشكلات بحوث تربوية.

3- بيان المعايير المستخدمة لتقييم مشكلة بحث ما.

4- تقييم مشكلة ما من أجل البحث وذلك باستخدام المعايير المقبولة.

5- اختيار إحدى المشكلات العامة في حقل اهتمام أحد الأشخاص في التربية، ثم وضعها بشكل محدد وجاهز للبحث التجريبي.

6- تعريف اصطلاحات مثل: المجتمع الإحصائي، والمتغيرات كما تستخدم في أي بحث من البحوث.

7- تشخيص المجتمع الإحصائي والمتغيرات في إحدى الدراسات المفترضة.

يبتدئ البحث النظامي بإحدى المشكلات. وقد تحدث جون ديوى (1933) عن الخطوة الأولى في الطريقة العلمية قائلا بأنها الإدراك بوجود صعوبة ملموسة، وعقبة أو مشكلة تحير الباحث.

إن اختيار وصياغة إحدى المشكلات هي أحد الأوجه المهمة للقيام ببحث في أي حقل من الحقول. والباحثون المبتدئون غالبا ما يندهشون عندما يجدون بأن هذه الخطوة الابتدائية غالبا ما تبتلع جزءا كبيرا من الوقت الكلي للقيام بمشروع البحث. وليس هناك وسيلة للقيام ببحث ما لم تكن هناك مشكلة يتم إدراكها والتفكير فيها وصياغتها بطريقة مفيدة.

ويتحتم على الباحث أولاً أن يحدد المجال العام للمشكلة. كما يتوجب على الباحث أن تكون له المعرفة أو الخبرة في ذلك المجال من أجل قيامه بتوجيه أسئلة يمكن إجابتها من خلال البحث. فنحن غالبا ما نسمع طلبة في مقررات دراسية صعبة يقولون: "إنني لا اعرف بما فيه الكفاية لطرح أسئلة". وبالمثل، ما لم تكن لباحث المعرفة أو الخبرة في أحد المجالات، فلن تكون لديه فكرة واضحة عن المعرفة الإضافية التي هناك حاجة إليها، أو كيفية الحصول عليها من خلال استقصاء تجريبي. والأكثر من ذلك، فإن مجال المشكلة المختارة للبحث يجب أن تكون ذات اهتمام عميق، أو يدور حولها فضول حقيقي. وبالضرورة، يجب أن يكون الاختيار شخصيا وإلا فإن الدافعية للاستمرار في البحث إلى النهاية قد يكون صعبا حشدها. وربما يهتم أحد معلمي المدارس الابتدائية بإيجاد طريقة أكثر كفاءة لتدريس القراءة. وربما يود أحد مدرسي الإنكليزية في إحدى المدارس الثانوية معرفة عما إذا كان استخدام (جهاز معالجة الكلمات) يحسن طريقة الكتابة لدى التلاميذ.

وحين يتم الاختيار، عندئذ يجري تضييق الموضوع العام، في عبارة محددة لمشكلة البحث. وأغلب الباحثين المبتدئين يجدون في مهمة صياغة مشكلة قابلة للبحث أو سؤال، أمراً صعباً. والصعوبة لا تعود إلى قلة المشكلات القابلة للبحث في التربية. وفي الحقيقة، هناك الكثير من الأسئلة التي تنتظر الإجابات حيث يجد الباحثون، عادة صعوبة في الاختيار بينها. وهناك إحدى المصاعب الشائعة في وجوب اختيار إحدى المشكلات وصياغة السؤال بصورة مبكرة بينما يكون فهم المبتدئ لكيفية القيام بالبحث محدوداً للغاية. بالإضافة إلى ذلك، فإن الالتباسات حول طبيعة مشكلات البحث، وعزل المشكلة، ومعايير قبولها، وكيفية حل المشكلة غالبا ما تبدو مربكة. فحتى الباحثون من ذوي الخبرة يجدون بأنه من الضروري القيام بعدة محاولات قبل التوصل إلى مشكلة بحثية تتصف بالمعايير المقبولة عموما. وربما يكون الانتقاء الأول أو الصياغة الأولى غير عملي ولا يستحق البحث عند التفحص الدقيق. والمهارة في القيام بالبحث تتعلق إلى درجة كبيرة، بالاختيارات الحكيمة لما يمكن استقصاؤه. والمهارة تستغرق وقتا وجهودا متكررة لتطويرها، لكنها يمكن أن تتطور لدى المبتدئ الراغب في ذلك.

ويجب أن تشير عبارة المشكلة، بوضوح، إلى ما يجب بحثه. وقد تكون الصياغة للمشكلة في جملة خبرية أو سؤال.

وفي كلتا الحالتين، يتحتم على الصياغة، الإشارة إلى المتغيرات ذات الأهمية والعلاقة المحددة بين المتغيرات التي سوف تجري دراستها. مثلاً، فإن المشكلة البحثية لمدرس المدرسة الثانوية

المذكورة آنفا يمكن أن تكون كالآتي: (المشكلة التي ستحري دراستها في هذا البحث هي تأثير استخدام (معالج الكلمات) على كتابة طلبة الإنكليزية في المدرسة الثانوية). أو، (ما هو تأثير استخدام معالج الكلمات على كتابة طلبة الإنكليزية؟ ويفضل العديد من الباحثين استخدام صيغة السؤال إذ إنها كما يبدو توجه المرء مباشرة نحو مهمة إيجاد الجواب، لكن كلا الصيغتين مقبولتان.

وليس محتملاً كما يبدو، بأنه إذا ما جرى انتقاء مجال المشكلة وصياغة المشكلة بسؤال أو عبارة بوضوح، أن تكون إحدى اصعب أوجه عملية البحث قد تم إنجازها.

طبيعة المشكلات THE NATURE OF PROBLEMS

بالرغم من وجود نماذج مختلفة من مشكلات البحث في مجال التربية غير أن كلها تتضمن تنقيباً عن المعرفة في الميدان. فالمشكلة توضح ما يود المرء معرفته كما تقوم بتحديد طريقة للشروع في دراستها للحصول على المعرفة.

ويقوم البحث التجريبي بتوجيه أسئلة تهتم بالعلاقات بين المتغيرات التي يمكن للباحث أن يتحكم بها. وفي البحث التجريبي يقوم الباحث بتفعيل المتغير المستقل (وهو السبب) بينما يقوم بالتحكم في جميع المتغيرات الأخرى، كما يسعى لملاحظة أية تغيرات لاحقة في المتغير التابع (وهذا هو الأثر). ويتيح التحكم خلال أية تجربة للباحث حذف أية تفسيرات بديلة للنتائج التي تمت ملاحظتها. فإذا ما رغب المرء ببحث طريقة التعليم على تعلم عمليات الجمع لأطفال المدرسة الابتدائية، فيتوجب على ذلك الشخص تفعيل طريقة التعليم بينما يتحكم / يضبط كل العوامل الأخرى التي قد تسهم في تعلم هذه العمليات، ومن ثم يلاحظ تأثير الطريقة على التعلم.

ويسعى البحث العلمي – المقارن أيضاً نحو تحديد العلاقات بين اثنين أو أكثر من المتغيرات، لكنه في هذا النوع من البحث يتعذر تفعيل المتغير المستقل. فالمتغير المستقل قد حدث فعلا: لهذا السبب فإن الباحث لا يتمكن من تقرير الكيفية التي يطبق بها المتغير المستقل. وإذا ما أراد باحث بحث / استقصاء تأثير طلاق الأبوين على أداء الأطفال، فيتحتم على الباحث مقارنة أداء المجموعات التي كانت موجودة قبلا: أحد الأطفال من أبوين منفصلين بالطلاق، وطفل ثان من أبوين يعيشان معا.

وهناك أنواع أخرى من البحوث الكمية، كالمسحية والارتباطية تطرح أسئلة عن طبيعة، ومدى تأثير، وتوزيع المتغيرات التربوية و/ أو العلاقات بين هذه المتغيرات. وليست ثمة محاولة لتفعيل المتغيرات – هناك، فقط، أوصاف للمتغيرات ولعلاقاتها كما تحدث بصورة طبيعية. واحد الأمثلة على هذا النموذج من مشكلة البحث هو السؤال التالي: ماذا يعرف طلبة السنة النهائية في المدارس الثانوية عن العملية التي من خلالها تقوم الأحزاب السياسية باختيار مرشحي رئاسة

الجمهورية؟ أو، ما هي العلاقة بين عدد سنوات دراسة الفرنسية في المدرسة الثانوية وتعيين مستوى المساق في الجامعة؟

وإذا ما رغب الباحث في بحث معنى التربية بالنسبة إلى العائلات الأمريكية ذات الأصل المكسيكي، و/ أو ما هو دور العائلة في تعليم الأطفال فسيستخدم المنهج النوعي. وسيقوم الباحث بدراسة عدد قليل من العائلات خلال مدة مطولة وسيجمع بيانات من خلال الملاحظات والمقابلات.

مصادر المشكلات THE SOURCES OF PROBLEMS

إن أول سؤال يستفسر عنه أغلب الطلبة هو: كيف أجد مشكلة بحثية؟ وبالرغم من عدم وجود قواعد ثابتة لتحديد مشكلة ما، إلا أن هناك اقتراحات معينة قد وجد بأنها مفيدة. فهنالك ثلاثة مصادر للمشكلات هي الخبرة، والاستنتاجات من النظريات وأدبيات البحوث(*) ذات الصلة.

الخبـــرة Experience

إن أكثر المصادر خصوبة، بالنسبة للباحثين المبتدئين، هي خبراتهم الذاتية كممارسين تربويين. فيجب اتخاذ القرارات يوميا بخصوص التأثيرات المحتملة للممارسات التعليمية على سلوك التلميذ. وإذا ما كان لهذه القرارات أن تكون صائبة، فيتحتم على التربويين القيام باستقصاء ناقد لصدق افتراضاتهم المختصة بالعلاقة بين الخبرات التعليمية وبين التغير في التلميذ.

هنالك قرارات يجب اتخاذها حول طرق التدريس. وبالتأكيد فإن طرق التدريس حساسة، وتحتاج لاستقصاء علمي. والطريقة العلمية في الممارسات التدريسية تؤكد بأن القرارات حول الكيفية التي تتم بها الأمور في التربية يجب أن تستند إلى دليل تجريبي وليس إلى شعور حدسي، أو انطباعات، أو معتقدات خاصة. فمثلاً، ربما يتساءل معلمو الابتدائية عن تأثير أساليبهم في تدريس القراءة. وقد يرغبون في تقييم طرقهم التدريسية العادية أو طريقة من الطرق المتعددة المشهورة لكي يقرروا أكثر الطرق تأثيراً. وربما يتساءل معلمو الثانوية فيما إذا كانت طريقة مناقشة المشكلة أو طريقة المحاضرة ذات تأثير أكبر في تدريس علم الاقتصاد في المدرسة الثانوية. وربما يرغب مدرسو الأحياء بحث تأثير المحاكاة المستندة إلى الحاسوب على تطوير مهارات الطلبة في حل المشكلات. فهل محاكاة الحاسوب أكثر فعالية من الطرق البديلة الأخرى؟

(*) يشير المصطلح "literature" بشكل عام إلى الأدب من صناعة ونتاج، وفضلنا استخدام "أدبيات" أو "أدبيات البحوث" في حقل منهج البحث العلمي للإشارة إلى البحوث والكتابات المتعلقة بهذا الحقل (المراجع).

إن ملاحظة علاقات معينة لا توجد لها تفسيرات مرضية هي أيضاً مصدر آخر للمشكلات التي يمكن البحث فيها. فقد يلاحظ أحد المعلمين زيادة في علامات القلق الظاهرة عند الطلبة في أوقات خاصة. ومن أجل البحث في ذلك، بوسع المعلم بناء تفسيرات متنوعة مؤقتة عن مصدر ذلك القلق ثم يبتدئ باختبارها تجريبياً. وقد لا يجد هذا البحث حلاً فورياً للمشكلة، لكنه يسهم قليلاً في عملية فهم أسباب القلق لدى تلاميذ الصف.

وبالمثل، هنالك قرارات يجب اتخاذها إزاء ممارسات قد أصبحت روتينية في حياة المدرسة وفي بعض الحالات كانت مرتكزة أساسا على التقاليد أو المرجعية، وتتمتع بإسناد قليل أو بدون إسناد يذكر من البحث العلمي. فلماذا لا يجري تقويم بعض من هذه الممارسات - مثلا: برنامج الاختبار السنوي؟ وهل هناك اختبارات بديلة تكون أكثر صدقاً بالنسبة للغاية ذاتها من تلك المستخدمة الآن؟

يمكن للخبرات اليومية للمربين إنتاج مشكلات تستحق البحث، وفي الواقع، فإن أغلب الأفكار البحثية التي طورها باحثون تربويون مبتدئون تبدو كأنها ناشئة من خبراتهم الذاتية. ولربما يكون لديهم مشاعر حدسية حول علاقات جديدة أو طرق بديلة لإنجاز أهداف معينة، وهكذا يصلون من خلال عملية حدسية إلى أفكار للبحث. وتلك الدراسات ستكون عموما إحدى النماذج المؤدية إلى حل لمشكلة آنية، ومع ذلك، في بعض الأحيان، تكون هذه المشكلات أكثر ملاءمة واكثر معنى بالنسبة للباحثين المبتدئين من تلك التي تم الوصول إليها من خلال استنتاجات منطقية من النظريات.

وبالإضافة إلى ذلك، فإن الدراسات المشتقة من خبرات المعلمين في الصف قد تساهم على نحو متميز في تحسين الممارسات التربوية.

وحاليا، ثمة حاجة كبيرة للبحوث التي تركز على كيفية تحسين الممارسات التربوية. فالمدارس الأمريكية تواجه عددا من المشكلات الحرجة، مثل التحاق أعداد متزايدة من طلبة ذوي مستويات متدنية ضمن الاتجاه التربوي السائد (Pallas, Natriello and McDill, 1989)، وتهيئة التلاميذ لميدان العمل حيث تتزايد حدة المنافسة بالإضافة إلى التقدم التكنولوجي (وزارة العمل الأمريكية/1991). ويمكن لنتائج الدراسات البحثية للممارسين أن تلعب دوراً مهماً في حل هذه المشكلات، وفي تحسين مدارسنا. وقد جاء في تقرير حديث للأكاديمية الوطنية التربوية (ص 1991/19): "بأن الجهود الحالية لتطبيق الإصلاحات المدرسية ذات القاعدة العريضة سيصيبها الإخفاق إذا لم تتوفر بحوث مناسبة لإرشاد وجهة التغير".

الاستنتاجات من النظرية Deductions From Theory

يمكن للاستنتاجات التي يتم اشتقاقها من نظريات سلوكية وتربوية مختلفة والتي يألفها الباحث أن تكون مصدرا ممتازاً للمشكلات. وتتضمن النظريات مبادئ عامة، حيث تطبيقها على مسائل

تربوية معنية، يبقى افتراضياً حتى يجري دعمها بالتجربة العلمية. ومن خلال البحوث، فحسب، بوسع المرء تحديد إمكانه ترجمة التعميمات المتضمنة في النظريات إلى توصيات محددة لتحسين الممارسات التربوية.

وبوسع الباحث توليد فرضيات، من النظرية، تبين الاستنتاجات المتوقعة في حالة عملية خاصة، وهذا معناه، أن الباحث يسأل عن العلاقات الموجودة بين المتغيرات التي سوف تجري ملاحظتها إذا ما قامت النظرية بصورة صحيحة بتلخيص الموقف. وبعد ذلك يقوم الباحث باختبار نظامي ليتأكد فيما إذا كانت البيانات التجريبية تدعم الفرضية ومن ثم النظرية.

ولا توفر النظرية، فقط أفكاراً كثيرة للمشكلات البحثية، بل تقوم كذلك بربط الأفكار بالمعرفة الراهنة. وبمقدور الباحث قراءة ملخصات نظريات في المجلات أو في الكتب المقررة، أو بالعودة إلى النص الأصلي للنظرية ذاتها. ويمكن للباحث بعدئذ، باستناده إلى النظرية، إجراء استنتاجات أو توقعات: إذا ما كان (هذا) صحيحاً، فلابد توقع ملاحظة (هذا). وربما يحاول الباحث أيضاً رؤية ما إذا كانت ستظل النظرية قائمة أو يمكن تطبيقها في ظروف تختلف عن الحالة التي جرى تطويرها من خلالها.

هنالك نظريات إدراكية، ونظريات لإثارة الدوافع، ونظريات تطورية/نمائية، ونظريات السمات، والكثير غيرها التي قدر يجري اختبار صدقها، ومجالها، وتطبيقاتها العملية بما يفيد المواقف التربوية. وقد تكون نظرية التعزيز، على وجه الخصوص، نقطة انطلاق مفيدة لبحث ما يجري في غرفة الصف. دعنا نفكر في مضامين الاختبار الصفي التي يمكن استنتاجها من مسلمة واحدة فقط من إحدى نظريات التعزيز – ونعني بذلك أن تعزيز الاستجابات يقود إلى زيادة في معدل الاستجابة وقوتها. ونحن نعلم بأن هذه النظرية قد أثارت كمية كبيرة من البحث، لكنه لا يزال الكثير من الاستنتاجات التي يجب إجراؤها واختبارها تحت ظروف صفية. فمثلاً، لم تكن هناك بحوث كافية بخصوص تأثير نقص التعزيز الظاهر أو عدمه على الاستجابات الصحيحة للطلبة في مواقف صفية. فنحن نعلم من دراسات مختبرية تجريبية مع الحيوانات بأن كل كبح للتعزيز يقلل احتمال تلك الاستجابة ويؤدي في نهاية الأمر إلى انطفائها. فهل يمكن نقل هذه النتيجة بتنبؤ احتمالي إلى غرفة الصف؟ وبكلمات أخرى هل يمكن للمعلم من الافتراض بأن الاستجابات الصحيحة للتلاميذ التي لم يجر تعزيزها ظاهريا سوف تضعف وتنطفئ بعدئذ؟ إنا في الوقت الراهن لا نملك بحوثاً صفية كافية من أجل اختبار تطبيقات هذا المبدأ.

ومن بين النظريات الأخرى التي كانت مصادر مفيدة لفرضيات خاصة بالبحوث الصفية: نظرية (Wittrock, 1974) للتعلم التوليدي، ونظرية التوقع (Resental & Jacobson, 1968). وطبقا للنظرية التوليدية، فالتعلم الجديد يحدث عندما يولد المتعلمون، بنشاط، علاقات بين ما فهموه سابقا مع ما سيتعلمونه. وقد افترض كل من (Kourilsky & Wittrock, 1992) بأن طلبة الثانوية الذين تعلموا استخدام أساليب الاستيعاب التوليدي في مقررات الاقتصاد سوف

يتعلمون أكثر وسيكونون أقل عرضة للخطأ من الطلبة الذين لم يتعلموا بهذه الطريقة. ونتائج التجربة ساندت هذه الفرضيات. واستنتج الباحثون بأن استخدام أساليب التعلم التوليدي يسهل التعلم بدرجة كبيرة.

والبحوث المبكرة المرتكزة على نظرية التوقع ساندت على العموم المقدمة المنطقية للنظرية التي تقول بأن توقعات المعلم تميل لان يكون لها تأثير على إنجاز التلميذ، وهذا معناه، بأن التوقعات العالية تميل لزيادة تحصيل التلميذ، والتوقعات المتدنية تميل لتخفيض تحصيل التلميذ. ويشير مضمون هذه النظرية في الممارسة إلى ضرورة احتفاظ المعلم بتوقعات عالية إذ ستكون نتيجته تحصيل أعلى للتلاميذ. والبحوث التي أنجزت مؤخراً (Jussim, 1989; Goldberg, 1992) قد وجدت بأن ظاهرة التوقعات ليست دائما دقيقة وقابلة للتنبؤ كما كان الظن في السابق، ولذلك دعت إلى بحوث أخرى.

إن طريقة معالجة المشكلات البحثية، المستندة إلى نظرية، ينتج عنها دراسات يمكن دمجها سوية بسهولة إذ أنها جميعا ترتكز على نظرية مشتركة. إن بحثا ذا علاقة متبادلة كهذا يكون منتجا وخصوصا كوسيلة لتوسيع المعرفة في نطاق معين.

الأدب ذو الصلة Related Literature

المصدر الثمين الآخر للمشكلات البحثية هو الأدب/الأدبيات في المجال الذاتي لاهتمام الفرد. ونتمكن عند قراءتنا لبحوث سابقة من الاطلاع على نماذج من مسائل بحثية، وعلى الطريقة التي سار عليها البحث. وثمة طرق عدة، تتيح لها مراجعة البحوث السابقة، صياغة مشكلات بحث جديدة.

1- تساعد مراجعة البحوث السابقة في صياغة أسئلة بحثية تعتبر الخطوة المنطقية التالية من استقصاءات سابقة. وغالبا ما تقود النتائج من أحد البحوث إلى أسئلة جديدة. وفي بعض الحالات، فإن الباحثين في الفصل الختامي من بحثهم يصفون أسئلة جديدة تكون قد برزت. وهناك طريقة منتجة لتوسيع الدراسات تتمثل في تقديم متغيرات جديدة من أجل إخضاعها لتحكم إضافي لاكتشاف تأثيرات التفاعل بين المتغيرات. وكثير من الدراسات المتعددة المتغيرات هي امتدادات لبحوث سابقة ذات متغير واحد.

2- وقد تثير مراجعة البحوث السابقة أحد الباحثين لرؤية ما إذا كان بالوسع تكييف الأساليب المستخدمة لحل مشكلات أخرى، أو ما إذا كان في الإمكان إجراء دراسة مماثلة في حقل مختلف أو في موضوع مختلف مع مجموعات مختلفة من الأفراد. فمثلا، يقرأ شخص ما إحدى الدراسات الباحثة في تأثير طريقة متعددة الوسائل في تدريس الكيمياء. وقد يمكن إجراء دراسة مماثلة في علم الأحياء أو في موضوع آخر. أو ربما تستخدم إحدى الدراسات الخاصة بطلبة المرحلة الثانوية كدليل لمعلم في المرحلة الابتدائية يهتم بتحديد ما إذا كانت العلاقات ذاتها

بين المتغيرات تبقى كما هي على مستوى الدراسة الابتدائية. وربما تجري إعادة دراسة، فعالية استخدام الحاسوب في تعليم القراءة لمجموعة منتظمة من تلاميذ السنة الأولى، على أفراد من الطلبة يعانون من عجز تعلمي.

3- قد توحي مراجعة بحوث سابقة بالرغبة في التكرار/الإعادة لتأكيد استنتاجات سابقة. وإحدى المميزات الأساسية للدراسة البحثية العلمية هي وجوب قابليتها للتكرار، كي تكون نتائج التجربة ممكنة التحقيق. إن تكرار الدراسة مع التغيير أو بدونه قد يكون أحد الأنشطة المفيدة لباحث ناشئ. إن إعادة دراسة ما يزيد مدى تعميم نتائج البحث، كما إنها توفر دليلا إضافيا لصدق النتائج. وفي كثير من التجارب التربوية لا يكون ممكناً اختيار أفراد عشوائياً، بل ينبغي استخدام المجموعات الصفية كما هي منظمة مسبقاً. وبالطبع فإن ذلك يحدد مدى تعميم نتائج البحث. وفي الوقت الذي تجري فيه تكرار التجارب في أوقات مختلفة وفي أماكن مختلفة تكون العلاقات المتوقعة مدعومة في كل دراسة، تزداد الثقة في الصدق العلمي للنتائج. إن مجرد إعادة الدراسات الأخرى وتكرارها لا يشكل أكبر الأنشطة البحثية تحديا، لكن هنالك، غالباً، حاجة في المشكلات التربوية لتأكيد النتائج وتوسيعها.

وفي أغلب الحالات فإن تكرار التجارب لدراسات سابقة لا تكون دقيقة. فيتم إبراز الاختلاف لتوضيح أحد وجوه النتائج، أو لاختبار مدى إمكانية تعميم نتائج البحث، أو لبحث عوامل لم تكن موجودة في الدراسة الأصلية. فمثلا، لقد جرت عدة تكرارات لدراسات بياجيه 1932 بخصوص تطور الأحكام الأخلاقية عند الأطفال وذلك في أقطار أخرى. وقد استخدمت هذه الدراسات الطريقة الأساسية لبياجيه، لكنها قامت ببحث تطور الأحكام الأخلاقية في أطفال جاؤا من طبقات اجتماعية- اقتصادية مختلفة، ولدى أطفال ذوي فئة عمرية واحدة لكنهم كانوا يختلفون في مستوى الذكاء، وعند أطفال مختلفين في مدى مشاركتهم في مجموعات فئاتهم العمرية، وفي أطفال يختلفون في طبيعة نظام أولياء أمورهم، وعند الأولاد والبنات على السواء. ومؤخرا استعمل باحثون آخرون تقنيات تختلف عن تقنيات بياجيه وذلك في محاولات لإثبات نتائج بحثه واستنتاجاته. وعموما، فإن الكمية الكبيرة من البحوث المنبثقة من استقصاءات بياجيه قد دعمت استنتاجاته الأصلية. وهكذا، إذا عالجت دراسة بحثية منفردة، مشكلة مهمة، وكانت نتائج بحثها مثيرة، يمكن لها أن تكون مصدر الهام لدراسات كثيرة أخرى.

4- قد تطرح دراسات بحثية سابقة سؤالا حول قابلية تطبيق نتائج بحوث تلك الدراسات في ثقافات أخرى. فنتائج بحوث جرت في إحدى الثقافات يتعذر تطبيقها أتوماتيكيا على ثقافات أخرى. وهذا هو السبب الذي دعا إلى وضع تأكيد كبير مؤخراً خلال العقود الحالية على بحوث ثقافات متعددة. وفي مجالات مثل علم نفس الأطفال والمراهقين، والتعلم الاجتماعي، وتطور الإدراك واللغة ودوافع الإنجاز وتطور الشخصية، والممارسات التربوية، كما أن هناك عدة أمثلة بالإمكان إيجادها في بحوث الثقافات المتعددة. فالعدد المتعاظم للطلبة غير

الأمريكيين في الجامعات الأمريكية قد زاد الاهتمام بهذا النوع من البحوث كما يسر تجميع بيانات الثقافات المتعددة. وقد قدم أولئك الطلبة عدداً من أطروحات الماجستير والدكتوراه من هذا النمط، وهذه تمثل قاعدة لدراسات قادمة في الثقافات المتعددة.

5- قد تسفر مراجعة بحوث سابقة عن اكتشاف عدم ثبات وتناقضات أو عدم رضى فيما يخص التصور العقلي، والمنهجية ووسائل القياس، والتحليل الإحصائي المستخدم. وغالباً، ما يتمكن الباحثون من إيجاد شيء لإدخال تحسينات على بحث سابق ما. فمثلا، تحدى كل من (Hartshorne & May, 1928) النظرية القائلة بأن النـزاهة سمة متحدة بالشخصية. فقد بحثا سلوك عدة آلاف من تلاميذ المدارس أثناء وجودهم في مواقف مغرية متنوعة. وقد استنتجا من الارتباطات المتدنية بين قياسات الإغراء، بأن النزاهة لم تكن سمة داخلية عامة، لكنها بالأحرى كانت محددة ومتأثرة بالموقف الذي وضع به الأفراد.

وقد أعاد (Burton, 1963) هذه الدراسة لأنه شك في تأكيد (هارتشورن وماي) على عوامل الظروف في السلوك النزيه. ولقد استخدم بيرتن طريقة (التحليل العاملي) ووجد دليلا على ثبات داخلية الفرد خلال المهمات التي عرضت على الأطفال. وقد استنتج انه ربما كانت هناك سمة ضمنية للنزاهة تظهر عند الفرد أثناء مقاومته لمواقف إغرائية، لكنه اتفق مع (هارتشورن وماي) في رفض صيغة (الكل أو لا شيء) فيما يتعلق بسجية الفرد.

وأعاد (Hunt, 1965) دراسة بيرتن، لكنه استخدم (تحليل التباين) بدلاً من التحليل العاملي كأسلوب إحصائي. وقد أشار (هنت) إلى أن الاختلافات الشخصية، والمواقف، والتفاعل بين الأشخاص والمواقف لابد من اعتبارها كمصادر للتغير في السلوك النـزيه، كما وجد دعماً لفرضية التفاعل.

وقام (Nelsen, Grinder & Muttere, 1969) بإعادة وتوسيع جميع الدراسات الآنفة. وقد قارنوا الطرائق المنهجية البديلة للمشكلة، واستنتجوا بأن السلوك الإغرائي يكون فقط ثابتاً باعتدال في تشكيلة متنوعة من المهمات. وهكذا، فإن نتائج بحوثهم تتفق مع الدراسة المبكرة التي أجراها كل من (هانت و ماى) بالرغم من اختلاف الزمن لأكثر من أربعين عاماً في فترات جمع البيانات لاثنتين من الدراسات، وبالرغم كذلك من اختلافات المجتمعات الإحصائية، والمهمات، والإجراءات المنهجية. وهنالك مثال آخر على البحث الذي أثارته دراسات سابقة، وهو البحث الذي يستقصي تأثيرات تجميع الطلبة حسب القدرات على تحصيل الطلبة. ومنذ السنوات الأولى للقرن العشرين جرت مئات الدراسات على هذه المشكلة. وبالرغم من ذلك لا تزال المشكلة موضوعا للبحث (Gamoran, 1993; George, 1993; Schmidt, 1993).

وغالباً، لا تتمتع نتائج البحث في موضوع معين بالثبات. وهذه هي الحالة في البحث الحالي حول تأثير استخدام (معالج الكلمات) في تدريس الكتابة. وتفيد بعض الدراسات

المنشورة بأن هنالك تأثيرات إيجابية، بينما تفيد أخرى بأن التأثيرات مختلطة. وهكذا، فإن استخدام (معالج الكلمات) في تدريس الكتابة للتلاميذ من المدرسة الابتدائية وحتى الكلية، موضوع جدير ببحوث أخرى.

وقراءة المعرفة المنظمة في حقل معين، قد تجعل الفرد على وعي بوجود فجوات معرفية. ويمكن إجراء تخطيط للبحث بشكل يساعد على ردم تلك الفجوات المعرفية، والوصول إلى معرفة يمكن الاعتماد عليها. وبالنسبة للطلبة، فإن النقاشات مع أساتذتهم في الحقل العام لدراستهم قد يكون مفيداً بمثل هذه الخطوة.

وينبغي أن يكون الطلبة قادرين على اكتشاف مشكلات عدة محتمل لها أن تكون صالحة للبحث وذلك بإجراء بعض التحليل النقدي للبحوث المنشورة في مجال دراستهم إضافة إلى قليل من الإبداع. إن فهم الأوجه التجريبية والنظرية لحقل المادة الدراسية يمكنهم من القراءة النقدية ومن اختيار مشكلة صالحة للبحث.

مصادر غير تربوية Noneducation Sources

بوسع خبراتنا وملاحظاتنا في العالم، عموماً، وكذلك أنشطتنا المهنية، أن توفر مصادر مفيدة لمشكلات بحثية. والنظريات أو الإجراءات التي تتم مواجهتها في حقول أخرى يمكن تكييفها للتطبيق في ميدان التربية. وغالباً ما تقودنا حركات انطلقت من خارج مهنتنا إلى مسارات جديدة في البحث. فالحركة النسائية قادتنا إلى دراسة التعميم النمطي للجنس في المواد التعليمية، وإلى تأثير المدارس على تعلم دور كل من الجنسين، وإلى اختلافهما في التحصيل وفي الشخصية، وغير ذلك. وأدت حركة الحقوق المدنية إلى دراسات كثيرة في تعليم أطفال الأقليات العرقية. أما انتشار مرض (نقص المناعة) في هذه البلاد فقد أثار الكثير من البحوث حول أفضل الإجراءات والمواد لتعريف اليافعين في المدارس بخطورة المرض واحسن الطرق لحماية أنفسهم منه. وجاء الإلهام لإجراء بحوث بالغة القيمة في مجال التربية من مصادر غير تربوية كهذه.

تقييم المشكلة EVALUATING THE PROBLEM

بعد أن يتم اختيار مؤقت لمشكلة بحث ما، يتوجب عندئذ تقييمها. ويتحتم على الباحث أن يكون واثقاً بأن مجال المشكلة مهم بما فيه الكفاية للقيام بعملية البحث، بالرغم من أن تقرير ذلك ليس بالأمر اليسير. فغالبا، ما يكون الحكم على قيمة مشكلة ما قضية قيم فردية ورأي شخصي. وبالرغم من هذا كله، فهنالك معايير معينة يجب استخدامها في هذه العملية الخاصة بتقييم أهمية مشكلة ما.

1- من الناحية المثالية، يجب أن يسهم حل المشكلة في مجموعة المعارف المنظمة للتربية.

ويتحتم على الباحث توضيح احتمال أن تقوم الدراسة بسد ثغرات في المعرفة القائمة، أو تساعد في حل حالة التقلبات في بحوث سابقة. وربما تطرأ تحسينات على الدراسة، في ضوء بعض الدراسات السابقة، بطريقة توفر معرفة أكثر موثوقية. ويتفق معظم العلماء بأن للمشكلات المتأصلة في النظريات، القدرة على تلبية هذا المعيار. وربما يقوم الباحث بالاستفسار فيما إذا كانت الدراسة ستقدم معرفة حول علاقات جديدة أو ستقوم بتكرار نتائج سابقة.

وهنالك احتمال أكبر بتركيز باحثين من ذوي الخبرة على النظريات، لكن الباحثين المبتدئين ملزمون بانتقاء مشكلات ذات تأثير على الممارسة التربوية. وقد قدم باحث تربوي محنك اسمه (Popham) هذه النصيحة: "من الأحسن قيام الباحثين المبتدئين بالنظر المدقق فيما يحدث حالياً في المدارس. وما هو أفضل، إدراك ما يحتمل حدوثه مستقبلا".(1999, p. 181) . ويتوجب أن يكون الباحثون قادرين على الإجابة على السؤال: (ثم ماذا؟) وذلك بخصوص دراستهم المقترحة. "فهل سيؤدي حل المشكلة إلى أي اختلاف في الممارسة التربوية؟ وهل يهتم مربون آخرون بنتائج البحث؟ وهل ستكون نتائج البحث مفيدة في موقف يتطلب اتخاذ قرار تربوي؟"

وفي محاولاتهم للعثور على مشكلة ما، يختار الطلبة في بعض الأحيان سؤالاً يتضمن علاقات سطحية إلى حد ما. مثلا، ما هي العلاقة بين شعبية الفرد مع أقرانه وسرعة القراءة؟ ويعد هذا سؤالاً قليل الأهمية بسبب ضآلة أهميته (إن وجدت) بالنسبة للممارسة التربوية، كما أن علاقته مع الدراسات الأخرى ضئيلة، ولا شأن له بالنسبة للنظريات. وأجوبة بعض الأسئلة واضحة. فالسؤال: "ما الارتباط بين الذكاء والتحصيل في القراءة؟" سوف يعتبر أيضاً قليل الأهمية إذ لدينا بيانات كافية حول هذه العلاقة، ومن ثم فإن بحوثاً أخرى يحتمل لها أن تكون غير ضرورية. فليس هناك حاجة لإعادة اختراع الدولاب، (أي لاختراع ما هو موجود – المراجع).

2- **يجب أن تؤدي المشكلة إلى مشكلات جديدة، وبالتالي إلى بحوث إضافية.** فإذا ما بدأ الباحثون بربط مسألة بحثهم بالمعرفة المنظمة، وأولوا بعض الاهتمام بنوع الدراسة التي ربما سوف تلزم منطقيا عن دراستهم، فمن المحتمل لهم، إلى حد كبير، تلبية هذا المعيار. فأثناء التوصل إلى إجابة لدراسة جيدة، غالبا ما يتولد العديد من أسئلة بحثية أخرى. وكان هذا صحيحاً في دراسات كانت تعالج نظرية التعزيز في غرفة الصف. وعلى العكس من ذلك، فإن الكثير من البحث المسحي يخفق في تلبية هذا المعيار الثاني.

وبالنسبة لهذا المعيار، فإنا نقترح بأن يقوم الطالب المبتدئ في البحث ببذل جهد دقيق أثناء اختيار إحدى المسائل بما يحتمل أن يقوم بتوسيعها أو بمتابعتها بعدئذ إما في رسالة ماجستير أو حتى رسالة دكتوراه.

3- **يجب أن تكون المشكلة قابلة للبحث.** وبالرغم من أن هذا المعيار واضح بذاته، إلا

أن الممارسة تفيد بأن هناك العديد من المشكلات التي لا تتضمن أسئلة يمكن إخضاعها لاستقصاء علمي. ولكي تكون المشكلة قابلة للبحث، ينبغي أن يكون ممكناً التصدي لها تجريبياً، وهذا معناه أن يكون بالوسع تجميع بيانات تتيح الإجابة على السؤال. وكثير من الأسئلة المثيرة للاهتمام في مجال التربية يتعذر إخضاعها للبحث التجريبي، غير أنه ينبغي بحثها من خلال استقصاء فلسفي. فأسئلة مثل: "هل من الصالح إدخال التربية الجنسية في المدارس الابتدائية؟ أو، هل ينبغي زيادة التدريب المهني في المدارس الثانوية؟ هي مسائل فلسفية لا يمكن الإجابة عليها بالبحث العلمي. وبالرغم من أنه تتعذر معالجة هذه الأسئلة، كما وردت، تجريبيا. إلا أنه يمكن إعادة صياغتها لتصبح أسئلة قابلة للبحث. فعلى سبيل المثال، يمكننا إعادة صياغة السؤال الأول أعلاه ليصبح كالآتي: "ما تأثير تدريس التربية الجنسية في المدارس الابتدائية على اتجاهات تلاميذ المدارس المتوسطة نحو العلاقة قبل الزواج؟ ومن الممكن تصميم إحدى الدراسات للحصول على معلومات حول هذا النمط من السؤال. وبالرغم من أن أسئلة فلسفية كهذه غير ملائمة للبحث العلمي، إلا أن المعلومات التي يوفرها البحث يمكن استخدامها في تطوير حلول لأسئلة فلسفية وأخلاقية، وهذا معناه، إن البيانات التي جمعت في البحث العلمي حول مشكلة ما يمكن أن تكون مفيدة للتربويين لدى اتخاذهم قرارات تتضمن الحقوق والقيم.

ويجب على الباحث أن يولي انتباها إلى تعريف وقياسات المتغيرات المتضمنة في السؤال. فمشكلة مثل: "ما هو تأثير التحولات في الأولويات الوطنية على مستقبل التربية الأمريكية؟" ليس مناسبا للبحث. وتعريف مصطلحات (تحولات في الأولويات الوطنية) و(مستقبل التربية الأمريكية) بطريقة تتيح قياسها سيكون مهمة صعبة.

4- يجب أن تكون مشكلة البحث ملائمة لذات الباحث: فقد تكون المشكلة ممتازة من حيث المعايير المذكورة، لكنها غير مناسبة للفرد. ونناقش فيما يأتي الجوانب الشخصية التي لابد من اعتبارها.

أولاً، يجب أن تكون المشكلة ذات أهمية حقيقية بالنسبة للباحث، بما يجعله متحمسا لها. ويتحتم أن يكون حل المشكلة مهم شخصيا بحيث يسهم بإضافة إلى معرفة الباحث، أو إلى تحسين أدائه كممارس تربوي. فما لم تكن المشكلة مفيدة ومهمة، فمن المشكوك فيه أن يكون الباحث راغبا في بذل الوقت والجهد للقيام بعمل متقن.

وفضلاً عن ذلك، يجب أن تكون المشكلة في مجال يعرفه الباحث وله فيه بعض الخبرة. فالباحث يحتاج أن تكون له معرفة بالنظريات القائمة، والمفاهيم، والحقائق القائمة ليتمكن من تشخيص مشكلة تستحق الاهتمام. والأكثر من ذلك، يجب على الباحث أن يفكر فيما إذا كانت لديه المهارات الضرورية التي قد يحتاجها للاستمرار في الدراسة وإكمالها. وربما يتوجب بناء الأدوات والمصادقة عليها أو تكون هناك حاجة إلى تحليلات إحصائية معقدة.

وهنالك اعتبار آخر وهو فيما إذا كان بحث المشكلة عمليا في الواقع القائم للباحث. كما ينبغي التأكد فيما إذا كانت البيانات الضرورية للإجابة على السؤال متوفرة أو قد تتوفر. وينبغي التأكد من توفر أفراد العينة اللازمة، أو إمكانية الوصول إلى سجلات المدرسة (المناسبة). وغالبا ما يعارض مديرو المدارس تنفيذ بحوث في مدارسهم. ولذلك، ما لم تكن تعمل في مدرسة ما في حينه، فمن المحتمل أن تجد نفسك دون أية وسائل لحل مشكلة البحث. ولقد وجد أحد المؤلفين أن عليه زيارة أربعة من الأنظمة المختلفة للمدارس قبل حصوله على الإذن لتنفيذ تجربة تربوية.

وكذلك، يتوجب عليك التفكير فيما إذا كانت المشكلة ممكنة للبحث ولإنهائها في الوقت المحدد لها. ويجب عليك أن لا تختار مشكلة كبيرة جدا أو تكون تضميناتها واسعة، كما ينبغي تخصيص الوقت الكافي لبناء الأدوات وتطبيقها، وتحليل البيانات، وكتابة التقرير.

5- **ينبغي أن تكون المشكلة ملائمة أخلاقيا.** وبكلمات أخرى، يجب أن لا تتسبب المشكلة المقصود بحثها بأذى بدني أو نفسي للعينات الإنسانية التي ستشارك في الدراسة. وسيكون الكثير لقوله عن أخلاقيات البحث في الفصل الرابع عشر.

عرض المشكلة STATING THE PROBLEM

بعد اختيار مشكلة البحث وتقرير أهميتها، لا تزال هنالك مهمة صياغة أو عرض المشكلة في قالب يتيح إخضاعها لعملية البحث. فالبيان الجيد للمشكلة يجب 1- أن يوضح بدقة ما يجب تقريره أو حله، و 2- حصر مجال الدراسة بسؤال محدد. فلا يمكننا أن نبالغ في تأكيد أهمية الوضوح والإيجاز في عرض المشكلة. وغالبا ما تكون لدى الباحثين المبتدئين فكرة عامة عن المشكلة لكنهم يواجهون صعوبة في صياغتها كسؤال بحثي قابل للتنفيذ. كما يجدون بأن أفكارهم العامة الأولية ليست محددة بما فيه الكفاية لأن تسمح بمعالجة تجريبية للمشكلة بالرغم من أن تلك الأفكار هي للفهم وللتواصل مع الآخرين. ولذلك فهم لا يتمكنون من إحراز أي تقدم ما لم يتمكنوا من صياغة سؤال ملموس يمكن إخضاعه للبحث.

وللتوضيح، افرض بأن باحثا مبتدئا يقول بأنه مهتم بدراسة تأثير مفردات المنهج الجديد في العلوم في المدارس الثانوية. وحسب عرض المشكلة، يمكن لأي فرد أن يفهم بصورة عامة ماذا يود أن يقوم به، كما يمكنه الحديث عنه بصورة عامة. لكنه يتحتم على الباحث تحديد المشكلة بوضوح شديد لإيجاد طريقة لبحثها.

وتعريف المفاهيم خطوة جوهرية لتحديد المصطلحات الواردة. فماذا يقصد بالمصطلحات: "تأثير/فاعلية، منهج العلوم، والمدارس الثانوية؟" إن التعاريف التي تحتاجها البحوث لا يوضحها القاموس، عادة. فمثلا، إن الفاعلية (التأثير) تعني: (تؤدي إلى نتيجة متوقعة أو مطلوبة)، وهذا التعريف يصف المفهوم البنائي (التأثير أو الفاعلية) عموما، إلا أنه ليس دقيقا بما فيه الكفاية

لأغراض البحث. ويحتاج الفرد إلى القدرة لتحديد مؤشر الفاعلية الذي سيستخدم في تحديد وجود أو انعدام الظاهرة المشار إليها بالمفهوم (التأثير /الفاعلية). وينطبق الأمر ذاته بصورة صحيحة على المصطلحات الأخرى. وبكلمات أخرى، يجب على المرء أن يعرّف متغيرات المشكلة إجرائيا. ومن أجل تعريف المفاهيم إجرائيا، يجب أن يحدد المرء، لتصور المفاهيم، نوعا ما من السلوك الظاهر أو الحدث الظاهر الذي يمكن ملاحظته مباشرة وقياسه عن طريق الباحث أو الآخرين. وكما ذكرنا في الفصل الأول، فإن التعريف الإجرائي يقوم بتعريف مفهوم ما بلغة الإجراءات أو العلميات التي تستخدم لقياس أو معالجة المفهوم.

وقد يختار الباحث في هذه الدراسة تعريف (الفاعلية) على أنها التحسن الحاصل في درجات أحد الاختبارات الخاصة بالتفكير الناقد أو أي اختبار مقنن للعلوم. أما المصطلح (منهج) فسيكون تعريفه بأنه مقرر علم الأحياء المعان بالحاسوب الخاص بطلبة الصف المتقدم في المدارس الثانوية. وقد يشير مصطلح (المدارس الثانوية) إلى مدارس المرحلة الأعلى التي لها مميزات محددة ومعينة مثل: الحجم، والنوع، وهكذا. والمشكلة الأصلية الآن ربما تكون: "ما هو تأثير مقرر علم الأحياء المعان بالحاسوب على استيعاب مفاهيم علم الأحياء لدى المبتدئين من الطلبة في دراسة هذا المقرر في الصفوف المتقدمة من المدارس الثانوية؟ فالتعاريف الإجرائية تساعد في تركيز مجال سؤال عام على متغيرات محددة وقابلة للملاحظة.

وبعد إتمام العمل ببعض من الوضوح والتركيز، يتمكن الباحث من تصميم دراسة تجريبية تقوم بمقارنة درجات اختبارات سابقة ولاحقة في استيعاب المفاهيم الأحيائية لتلاميذ يساعدهم الحاسوب في مقرر الأحياء وبين تلاميذ آخرين لديهم منهج بديل في علم الأحياء. ويتمكن الباحث الآن من الابتداء في جمع بعض الأدلة الموضوعية المتعلقة بمنهج معين في موقف معين بما سيلقي الضوء على السؤال الأصلي الأكثر عمومية.

وإضافة إلى ذلك، يجب على الباحث أن يعمل بجد من أجل إيجاد توازن بين التعميم والتخصص في عرضه للمشكلة. فإذا ما كان منطوق المشكلة واسعا وعاما أكثر مما ينبغي، عندئذ يواجه المرء مجالا غامضا دون إشارة واضحة إلى المسار الذي سيتبعه البحث. وعلى سبيل المثال، فإن سؤالا مثل: "ما تأثير التدريس المعان بالحاسوب على التحصيل في الحساب؟" هو سؤال عام جداً. ومن الأفضل هو الاستفسار: "ما تأثير استخدام برامج الجبر المعان بالحاسوب مقابل التدريس بالكتاب على التحصيل في الجبر لتلاميذ أذكياء في مدارس ثانوية؟ وهذا المنطوق يشير فورا إلى المجتمع الإحصائي، والمتغيرات المتضمنة، ونمط البيانات التي سوف يتم جمعها.

ومن الناحية الأخرى، يجب أن لا تكون المشكلة ضيقة المجال بحيث تصبح تافهة وبلا معنى، فالمشكلة يجب أن تكون واسعة بما فيه الكفاية لتكون ذات أهمية طبقا للمعايير التي نوقشت، وفي الوقت ذاته محددة بما فيه الكفاية لتكون عملية موقف خاص يهم الفرد.

وفي الوقت الذي تركز فيه مشكلات البحث عموما على العلاقة بين متغيرين أو أكثر، فهذا لا يعني بالضبط بأن الكلمات: "ما هي العلاقة بين.. وبين..؟" يجب أن تظهر في المنطوق. فقد يكون المنطوق بتلك الصيغة، أو قد تكون العلاقة ضمنية فحسب. ولا يتوجب بالضرورة عرض المشكلة بصيغة سؤال. وغالباً ما يقع الطلبة في الارتباك حول هذه النقطة. فمثلا، قارن المنطوق التقريري الآتي: "هذه الدراسة تتفحص التأثير الرئيسي لطريقتي القراءة (أ) و (ب) في تدريس بطيئي التعلم، مع السؤال الآتي: ما هو التأثير النسبي لطريقة القراءة (أ) بمقارنتها مع الطريقة (ب) في تدريس بطيئي التعلم؟ وكلتا الطريقتان تبحثان العلاقة بين متغيرات دون استخدام الكلمات الدقيقة، وكلتاهما طريقتان مقبولتان لعرض مشكلة البحث. وكما ذكرنا سابقا، فإن بعض المؤلفين يفضلون صيغة السؤال بسبب كونها مباشرة كما تبدو، نفسياً، وكأنها توجه الباحث نحو المهمة القائمة، ونقصد من ذلك، قيامه بتخطيط طريقة لإيجاد إجابة على ذلك السؤال. وفي بعض الحالات يجري عرض منطوق تقريري بالإضافة إلى سؤال، وربما تلي ذلك عبارة تقريرية عامة يعقبها سؤال أو عدة أسئلة محددة.

ويتوجب عليك كتابة مشكلة البحث بطريقة تجعل البحث في السؤال ممكنا. وتجنب القضايا الفلسفية، وكذلك الأسئلة القيمية أو أسئلة الرأي التي لا يمكن إجابتها بالبحوث العلمية.

وإذا ما تمت صياغة محتملة للسؤال، ينبغي عندئذ تقييم وضوحه وملاءمته. فتطوير مشكلة قابلة للتنفيذ، عملية تطورية تتضمن عدة محاولات لشحذ المفاهيم، وتحديد الإجراءات، والتفكير بطرق تجميع البيانات.

تحديد المجتمع الإحصائي والمتغيرات

IDENTIFYING POPULATION AND VARIABLES

الإستراتيجية الجيدة لتأطير مشكلة يتم الشعور بها – أو فكرة غامضة حول ما يريد شخص ما أن يبحثه– إلى مشكلة يمكن بحثها تتمثل في التفكير بلغة المجتمع الإحصائي والمتغيرات.

فمثلا، دعونا نفكر في السيدة بيرك، وهي مديرة مدرسة ابتدائية، والمسألة الصعبة التي تواجهها هي الآتي: هل للتدريس الفردي الذي يقوم به تلاميذ الصفوف العالية تأثير إيجابي على قارئين مستواهم دون المتوسط؟

من السهل جدا تحديد المجتمع الإحصائي، وهم أولئك الناس الذين يود المرء أن يعلم شيئا عنهم. فالمجتمع الإحصائي هم القارئون الذين مستواهم دون المتوسط. والقدرة على القراءة (القدرة القرائية) هي ليست متغيرا في هذه المسألة لأن جميع الأطفال المعنيين قد تم تشخيصهم وهم قارئون مستواهم دون المتوسط. وبعد قيام السيدة بيرك بهذا التحديد، عليها أن تتساءل

فيما إذا كان ذلك في الواقع، المجتمع الإحصائي الذي تريده أم لا. وربما ستقرر بأن (القارئين من ذوي المستوى دون المتوسط) شريحة عريضة، ويتحتم عليها الاقتصار على أعمار معينة. فتقوم باختيار قارئين من ذوي المستوى دون المتوسط الذين هم مجموعة الصف الثاني الابتدائي.

والآن هي متأهبة للبحث عن متغيرات في ما بقي من بيان منطوقها الأصلي. ويمكن اتخاذ (التدريس الخصوصي الفردي) كمتغير إما بتغيير نمط التدريس المستخدم، أو بتغيير زمن التدريس، أو بتلقي بعض الأطفال تدريسا خاصا، بينما لا يتعلم الآخرون بهذه الطريقة من الدراسة. وبالتالي تقرر السيدة بيرك بأن البديل الأخير يخص الشيء الذي تود أن تعرفه حقيقة، فتقوم بإعادة كتابة الجزء المطلوب من المسألة كالآتي: "هل تلقي مقدار معين من التدريس الخصوصي مقابل عدم التدريس الخصوصي ... ؟ وهكذا، يصبح التدريس الخصوصي هو المتغير المستقل إذ أنه سابق لتحصيل القراءة، ومديرة المدرسة تتوقع بأن التدريس سوف يكون له تأثير على تحصيل القراءة وهو المتغير التابع.

والآن يصبح واضحا بأن كلمة "التدريس الخصوصي" عامة أكثر مما ينبغي. وما لم يتلقى جميع أفراد العينة ذات النمط والكمية من التدريس الخصوصي، فإن نتائج الدراسة ستكون لا معنى لها. وهنا تقرر السيدة بيرك استعمال تمرين الكلمات الومضية/الخاطفة كنمط معين للتدريس الخصوصي مع تحديد 15 دقيقة في اليوم كمقدار للزمن.

ونتبين الآن أن عبارة "لها تأثير إيجابي على ..." غامضة تماماً إلى أن تتفحصها بلغة المتغير المستقل. هل لتمرين الكلمات الومضية تأثير على.. ماذا؟ إنها تدرك أن لها تأثيرا على تذكر الكلمات الومضية، لكنها تود دراسة تأثيراتها على الأوجه الأخرى من السلوك القرائي الذي يمكن ملاحظته: قراءة شفهية تعبيرية، قراءة صامتة، مشاعر إيجابية إزاء القراءة، عدد الكتب المقروءة، الاستيعاب، وهلم جرا. لكنها تخشى أن يصنف المعلمون من يتذكرون الكلمات بأنهم أكثر استيعابا، وبالتالي أكثر إيجابية نحو القراءة بينما يضعون من يكون تذكرهم للكلمات ضعيفاً في مرتبة أدنى، وفقا لقياس المتغيرات، مما هم عليه فعلاً. إنها بحاجة إلى متغير تابع لا يخضع لحكم المعلم، ومن ثم تقرر استخدام درجات القراءة في اختبار كاليفورنيا للتحصيل (CAT) كمتغير تابع.

فجملة بيرك المنقحة تصبح الآن: "هل هناك فروق في درجات CAT لمن مستواهم دون المتوسط من تلاميذ الصف الثاني الابتدائي، بين أولئك الذين يدربهم فردياً، لخمسة عشر دقيقة يوميا على تمرين الكلمات الومضية السريعة، تلاميذ صفوف أعلى، وبين أولئك الذين لم يتلقوا أية تمارين تدريبية؟ إن هذا السؤال يدلنا على الذين خصتهم السيدة بيرك في دراستها، وعما سوف يقدم إلى بعض التلاميذ دون أقرانهم الآخرين، وما تتوقع من تأثير معالجة متباينة. لاحظوا أيضاً بأن القرار القيمي (التأثير الإيجابي) قد أسقط من السؤال.

وغالبا ما يكن من المفيد اتباع هذا الإجراء بطريقة شكلية (منهجية) تشبه ذلك المتبع في

رسم تخطيط بياني لجملة ما. ويمكن للمرء الابتداء برسم خط عمودي، وكتابة كلمة (المجتمع الإحصائي) في اليمين، وكلمة (المتغيرات) في اليسار(*). ثم يجري تدوين هذه العوامل في الدراسة تحت الخط الأفقي، وحسب المثال أعلاه فإن الرسم التخطيطي سيكون كالآتي:

المجتمع الإحصائي	المتغيـــرات
القارئون من تلامذة الصف الثاني الذين مستواهم القرائي دون المتوسط	تدريب الكلمات الومضية لمدة (15) دقيقة يوميا يقوم به تلاميذ الصفوف الأعلى مقابل عدم وجود تدريب للكلمات الومضية (متغير مستقل) درجات القراءة في اختيار كاليفورنيا في الأداء CAT (متغير تابع)

دعونا الآن نأخذ مسألة أخرى: ما تأثير برنامج لما قبل المدرسة، يتلقاه بعضهم ولا يتلقاه آخرون – على الأداء في القراءة لدى تلاميذ الصف الأول الابتدائي؟

المجتمع الإحصائي	المتغيـــرات
تلاميذ الصف الأول الابتدائي	تلقي برنامج لفترة ما قبل المدرسة– وعدم تلقيه – (متغير مستقل) التحصيل القرائي (متغير تابع)

وهذا السؤال تام حيث أنه حدد المجتمع الإحصائي وكلا من المتغيرين، المستقل والتابع. وبما أن (برنامج ما قبل المدرسة) يسبق (الأداء في القراءة لدى تلاميذ الأول الابتدائي)، فالسابق يمكن تحديده كمتغير مستقل، واللاحق كمتغير تابع.

لنأخذ مثالاً آخر: هل التعليم المدرسي لقيادة السيارة مفيد؟ والسؤال كما هو وارد لا يحدد مجتمعا ولا متغيرات. فالباحث الذي يشرع بهذا السؤال قد يقرر مقارنة السائقين ذوي التاسعة عشر ربيعا الذين حازوا على تعليم مدرسي لقيادة السيارات في المرحلة الثانوية مع أولئك الذين لم يحوزوه. والآن أمامنا بيان مجتمع إحصائي ومتغير مستقل. ونتمكن الآن من توجيه انتباهنا نحو اختيار متغير تابع. ما تأثير التعليم المدرسي لقيادة السيارات مقابل عدمه لدى سائقي السيارات ذوي التسعة عشر ربيعا؟ لقد تقرر أن (نسبة الحوادث) سوف تكون متغيرا تابعا ملائما. والرسم التخطيطي أدناه يوضح ذلك:

المجتمع الإحصائي	المتغيـــرات
سائقو مركبات عمرهم 19 سنة	الحصول وعدم الحصول على التعليمي المدرسي لقيادة السيارات في المرحلة الثانوية (متغير مستقل). النسبة المئوية للحوادث (متغير تابع)

(*) كتب المؤلفون اليسار بدلا من اليمين وبالعكس، وقد أجرينا تعديلاً للانسجام مع طريقة الكتابة العربية من اليمين إلى اليسار. (المراجع) ،

ويمكن الآن بيان منطوق سؤال تام: هل يتعرض سائقو المركاب الذين عمرهم (19) سنة ولديهم تعليم ثانوي في السياقة، لنسبة حوادث أقل من نظرائهم الذين ليس لديهم تعليم ثانوي في السياقة؟

أما السؤال: "ما هي العلاقة بين مبدأ الجزمية (الدوغماتية) والاتجاهات السياسية بين طلبة السنة الأولى في الكليات؟" فهو يوضح نقطة أخرى.

انظر إلى هذا التخطيط البياني:

المجتمع الإحصائي	المتغيـــرات
طلبة السنة الأولى في الجامعة	مبدأ (الجزمية) مواقف سياسية

ويكون هذا السؤال متكاملا بوجود (المجتمع الإحصائي) واثنين من المتغيرات. لكننا لا نتمكن من تحديد المتغيرات كمستقل وتابع إذ لا يمكننا تقرير السابق منهما للآخر.

وإذا ما كانت هنالك دراسة تبحث في الأمر الراهن وليس في العلاقة بين المتغيرات، فقد تكون تامة بوجود متغير واحد.فمثلا، ربما يقوم شخص ما بدراسة آراء طلبة التخرج في الكلية بخصوص الإجازة القانونية للماريجوانا. وفي هذه الحالة يكون المجتمع الإحصائي هو طلبة في الكلية والمتغير الوحيد هو آراؤهم حول الموضوع.

وتستخدم طرق مختلفة للإجابة على الأنماط المختلفة من الأسئلة البحثية. ومتى ما يكون بوسع الباحث تفعيل (معالجة) أحد المتغيرات، فإن الطريقة التجريبية هي الطريقة الملائمة للاستخدام (انظر الفصل التاسع). فالمثال الأول في هذه الفقرة، المتضمن تدريب الكلمات الومضية مقابل عدم التدريب حسب درجات CAT في القراءة يمثل بحثا تجريبيا. والكثير من المتغيرات في التربية لا يمكن تفعيلها، وهكذا فإن الطريقة التجريبية غير ممكنة الاستخدام في البحث حسب هذه المتغيرات. والمشكلة البحثية المتضمنة مقارنة نسبة الحوادث للسائقين في عمر التاسعة عشر الذين يحصلون أو لا يحصلون على تدريب السياقة ستحتاج إلى الطريقة العلية – المقارنة (انظر الفصل العاشر). ولم يقم الباحث بتفعيل المتغير المستقل (تدريب السائق مقابل عدم التدريب).

وبالنسبة لبعض المشكلات البحثية فإن الطريقة الملائمة هي التجريبية أو العلية – المقارنة، وذلك يعتمد على كيفية تصميم الدراسة. والمثال البحثي أعلاه حول تأثير برامج ما قبل المدرسة على التحصيل في القراءة يمكن أن يستخدم أيا من الطريقتين، ويعتمد ذلك على إذا ما اختار الباحث تفعيل المتغير المستقل (برامج ما قبل المدرسة) أو اختار أفراد العينة الذين تدربوا أو لم يتدربوا في تلك البرامج.

وتختص بعض مشكلات البحث بوصف طبيعة أو مدى تأثير واحد أو أكثر من المتغيرات التربوية، أو العلاقة القائمة بين المتغيرات. والمثال المتضمن العلاقة بين مذهب (الجزمية) والاتجاهات السياسية بين طلبة السنة الأولى في الجامعة هو أحد الأمثلة على البحث الارتباطي. أما الدراسات التي تبحث عن آراء في قضايا مثل تشريع الماريجوانا، فهي تمثل بحوثا مسحية. كما إن المشكلات البحثية الأخرى، مثل حالة دراسة معهد اجتماعي معين، فإنها تحتاج إلى طرق نوعية (انظر الفصل الثالث عشر)

مقالة المجلة الدورية The Journal Article

إن تحديد المجتمع الإحصائي والمتغيرات المستقلة والتابعة يكوّن إطارا لتحليل التقارير البحثية المنشورة. ومقالة المجلة الدورية التي كتبها (March, 1989) التي تظهر في الصفحات التالية توضح العناصر التي توجد عادة في تقرير عن تجربة. فالمقدمة تبين الأساس المنطقي للدراسة كما تلخص التفكير والبحوث السابقة حول الموضوع. أما فقرة الطريقة فتحتوي على عدة عناصر: 1- العينة (عشرة صفوف من خمس جامعات) تمثل المجتمع الإحصائي ذا الأهمية (وهم طلبة الجامعات). 2- المتغير المستقل (اختبارات بيتية مقابل اختبارات صفية) قد جرى وصفه. وتعريفه إجرائياً. 3- جرى وصف المتغيرات التابعة (استجابات مسح يسأل الطلبة عن مدى دافعيتهم للدراسة والتعلم وذلك للاستعداد لكل اختبار، وكذلك درجات اختبار فجائي للاستيعاب وتطبيق المادة المقررة)، كما جرى تعريفها إجرائياً. 4- جرى وصف الإجراءات بتفصيل كاف لتمكين باحث آخر من إعادة التجربة ولكن بعينة مختلفة. أما فقرة النتائج فتقدم البيانات الناشئة عن التجربة، كما تقدم نتائج الاختبارات الإحصائية. وتقدم فقرة الخاتمة فإنها تقدم تفسير المؤلف لتلك النتائج.

مقارنـــة بين اختبارات
تجرى في الصف مقابل اختبارات تجرى في البيت

تمثل هذه الدراسة جهدا لتحديد ما إذا كان الاختبار البيتي أداة جيدة أو أفضل للتعلم مقارنة مع الاختبار التقليدي الصفي. ونتائج هذه التجربة إذ إن المعلمين يهملون الآن بصورة متزايدة الاختبارات الصفية (Gay and Gallagher, 1976). ويبدو أنه لا يوجد أساس تجريبي للاعتقاد بأن الاختبارات البيتية ذات تأثير. لذلك، يبدو بأن النزعة الحالية لاستبعاد الاختبارات الشكلية المكتوبة قد تعمل ضد واحد من أكبر أهداف التربية: وهو حفظ الطلبة للمفاهيم التي تعلموها.

وربما يعود القلق الكبير حول الاختبارات الصفية إلى كونها تتسبب في جزع مضني. وبالرغم من هذا كله، فإن البحث في هذا الإدعاء غير متماسك. فمثلاً، تبين أن لجزع

الموقف الاختباري تأثير سلبي (Sarason Davidson, Lighthall, Waite, and Ruebush, 1960)، وله تأثير إيجابي (Castaneda, and Palemo, 1974)، كما أنه لا تأثير له (Denny, Paterson, and Feldhusen, 1964) على الأداء في الاختبار.

وافترض بعض المؤلفين بأن مقدار القلق / الجزع هو الذي يؤثر في الأداء الاختباري، وتذهب (Anastasi, 1976) إلى أن للاختبار تأثيراً إيجابياً كلياً إذا ما أنجز بصورة ملائمة. وهي تقول بأن العلاقة بين الجزع والأداء في الاختبار ليست خطية، وأن بعض الجزع مفيد، بينما الجزع البالغ أمر مؤذ. وقد تم تصميم العديد من التطبيقات لإزالة الجزع في البيئة الاختبارية للصف، لكنها تشير إلى أن "تصرف الممتحن والإجراء المنظم جيداً والسلس للامتحان تساهم في السير نحو الهدف ذاته" (Anastasi, 1976).

وبالرغم من تعرض الطلبة لمستويات مختلفة من الجزع قبل الاختبارات، فإن فكرة إزالة الاختبارات لذلك السبب وحده قد لا تكون صالحة. فهدفنا يجب أن لا يكون تجنب الضغوط كلية، وهو شيء غير ممكن في عالمنا، بل أن نتعلم كيف ندرك استجابتنا النموذجية للضغوط ونحاول تعديل أسلوب مجابهته طبقا لذلك. (Selye, 1978).

ثمة ندرة في الأدبيات الخاصة بمقارنة الاختبارات البيتية بالصفية. وهنالك دليل على أن للاختبارات الشفهية والتحريرية تأثيراً، وأن أيا منهما أفضل من عدم وجود الاختبار أبداً (Calhoun, 1962). وتبين الدراسات كذلك بأن التحصيل الأكاديمي لطلبة المرحلة الجامعية الأولى هو أدنى بموجب نظام درجات ناجح/ راسب من نظام الدرجات المعياري، (Bain, Hales, and Rand, 1973; Golod, Reilly, Silbeman, and Lehr, 1971).

وفي سنة 1976 جرت دراسة (Gay and Gallagher, 1976) كان فيها صف كبير من طلبة المرحلة الجامعة الأولى تم توزيعها عشوائياً في ثلاث مجموعات. و أعطيت المجموعة الأولى تمارين بيتية على فترات. والمجموعة الثانية أعطيت اختبارات على فترات متقطعة. أما المجموعة الثالثة فقد أعطيت حرية الاختيار. ومن المفيد ملاحظة أنه لم يكن هناك أحد من المجموعة الثالثة قد اختار الاختبار. وما عدا هذه الاختلافات المذكورة أعلاه، فقد عومل جميع الطلبة بالتساوي بقدر الإمكان. وفي نهاية المقرر الدراسي أعطي الطلبة اختباراً مفاجئاً، ومتشابهاً في المجموعات الثلاثة جميعها. وأظهرت نتائج الاختبارات بأن المجموعة التي أعطيت في السابق اختبارات قد حصلت على درجات أعلى من الاثنتين الباقيتين بصورة بارزة.

وتختلف هذه الدراسة عن سابقتها بطرق عدة. أولاً، اشتملت على صفوف متنوعة في الآداب وفي الحقول السلوكية والتطبيقية للدراسة. وبناء على ذلك، من المحتمل تعميم النتائج على جمهور أوسع. ولقد حاولت الدراسة قياس مستويات معينة من التعلم كما عرفها التسلسل الهرمي للتعلم الإدراكي / المعرفي لبلوم (Bloom, Engelhart, Furst, Hill, Krathwohl, 1956). وأخيراً، فقد جرت محاولة لتشخيص العلّية بالنسبة إلى نتائج التجربة.

الطرق Methods

لقد تم انتقاء عشرة صفوف دراسية من خمس جامعات مختلفة كمشاركين. وقد كان (258) طالباً في العينة الكلية، التي اشتملت على صفوف من علم النفس، والإحصاء، والتربية، والبحث، والتاريخ. وفي الأيام الأولى للفصل الدراسي، أعطي كل مدرس مشارك ملخصاً شفهياً عن الإجراءات ومجموعة من الإرشادات لاتباعها حرفياً. وقد جرى إبلاغ الطلبة بأن نصفهم سيتلقى اختباراً بيتياً، وإن ذلك كان تجربة علمية، وإن نتائجها سوف لن تؤذي أحداً، وإن أي طالب غير مقتنع بدرجته سيكون أمامه الخيار للاختبار مرة أخرى. وكان من المتوقع بأن ذلك الإجراء سيكون له تأثير طبيعي على عامل القلق.

وقبل منتصف الفصل بأسبوعين، تم تقسيم الصفوف الدراسية عشوائياً إلى مجموعتين متساويتين: أ، ب. وفي كل صف من الصفوف اختير عشوائياً. أفراد من المجموعة (أ) وأعطوا اختباراً بيتياً على أن يتم إرجاعه بعد أسبوع. أما أفراد المجموعة (ب) فقد تم إعطاؤهم اختباراً صفيا مطابقا. وكانت جميع الاختبارات المستخدمة في هذه الدراسة من نمط الاختيار من متعدد. وقد انتهت المجموعة (أ) من الاختبار في ذات الوقت الذي بدأت فيه المجموعة (ب) اختبارها الصفي. وبعد مرور أسبوع أعطيت المجموعتان اختبارا فجائياً خلال فترة الدرس النظامية. وقد غطى الاختبار المادة التي اختبروا فيها في الأسبوع المنصرم. وبما أن أفراد المجموعتين كلتيهما قد قام بتدريسهم المدرسون ذاتهم وتحت الظروف ذاتها، فإن أفراد التجريبية الضابطة تعرضت لتعليم مكافئ حسب الإمكان في ظل تجربة نمط حجرة الدراسة.

لقد احتوى الاختبار الفجائي على خمسين سؤالاً في الاختيار من متعدد، خمسة وعشرون منها قاست المعرفة والباقية كانت خليطاً من أسئلة الاستيعاب والتطبيق في مادة المقرر الدراسي. واشتملت الأخيرة على المستويات الثلاثة الأولى من التعلم الإدراكي / المعرفي في الهرم التسلسلي لبلوم (Bloom et al., 1956). وقد ألحق مسح بالاختبار للمساعدة في تحديد مدى دافعية الطلبة نحو الدراسة والتعلم للاستعداد لكل نموذج من نماذج الاختبارات. وكانت نسبة الاستجابات هي 96.1% للاستبيان. وأخيراً، استخدم تحليل التباين ذي الطريقين (Two-way analysis of variance). ونظراً لوجود عشرة صفوف، فقد كان التصميم 10×2 (Ary and Jacobs, 1976).

النتائج Results

نالت المجموعة (ب) درجات دالة أعلى من المجموعة (أ) التي تلقت الاختبار البيتي في جميع الفئات من الاختبار الفجائي. وقد أظهر الأفراد، على نحو متميز، زيادة في مجمل التعلم والمعرفة والاستيعاب والتطبيق للمقرر الدراسي. وقد أوضح التحليل الإحصائي لكل صنف بأن ارتباط المتغير المستقل (نوع الاختبار) والمتغير التابع (التعلم) كان عالياً.

وفي المسح المرفق مع الاختبار الفجائي كان هناك سؤال يقول: ((لو كنت معلماً وأردت أن يتعلم تلاميذك فعلاً، فأي نوع من الاختبار ستعطيهم؟)). وقد كانت استجابة المشاركين: 25.8% يفضلون إعطاء الاختبار البيتي، بينما 59.1% فضلوا الاختبار الصفي، و 15.1% من الطلبة لم يكونوا متأكدين من الاختيار. ولقد بدوا وكأنهم يقولون بأن الاختبار الصفي كان أكثر تأثيراً وذلك بهامش يساوي (1:2).

ولقد تم تصميم سؤالين لتحديد اتجاهات المشاركين نحو الاستعداد لنموذج خاص من الاختبار. وطلب السؤال الأول من المشارك بأن يضع نفسه على تسلسل من واحد إلى عشرة مشيراً إلى مقدار الجهد الذي تم بذله في دارسة المادة التي اشتمل عليها اختبار المدرس قبل الاختبار الفجائي بأسبوع واحد. أما السؤال الثاني، وهو أيضاً ذو تسلسل من واحد إلى عشرة، فقد طلب مقدار الجهد الذي كان الطالب سيبذله فيما لو تقرر أن كان الاختبار الآخر (بيتيا أو صفياً).

وقد حصلت مجموعة (أ) على درجات، دالة، أقل في السؤال الأول. وهكذا، بالفعل، قال الطلبة بأنهم كانوا سيبذلون جهداً أشد إذا ما كان عليهم تلقى الاختبار الصفي. وبالعكس، فإن المجموعة (ب) قد حصلت على درجات، دالة، أقل بكثير في إجابتها على السؤال الثاني. ويمكن للشخص أن يفسر ذلك بالقول بأنهم كانوا سيبذلون جهداً أقل إذا ما كان عليهم تلقي الاختبار البيتي. وأحد الاستنتاجات الذي يمكن أن يفسر الاختلاف في التعلم بين المجموعات هو إن أولئك الطلبة الذين كان سيطبق عليهم الاختبار الصفي قد بذلوا جهداً أكثر حسب اعترافاتهم الذاتية.

النتيجة Conclusion

إن هذه البيانات تقدم دليلا لصالح الاختبارات الصفية مقابل تلك التي تجري في البيت لتشكيلة متنوعة من طلبة الكلية في أقسام: علم النفس والإحصاء والبحث، والتربية والتاريخ. وفي هذه التجربة فإن نوع الاختبار الذي استخدم كان له تأثير على مقدار تعلم الطلاب في المستويات الثلاثة الأولى في تسلسل (بلوم) الهرمي: المعرفة، والاستيعاب، والتطبيق. وبصورة واضحة فإن، الاختبار الصفي كأن الأحسن، والطلبة أقرّوا بذلك.

لكن – لأسباب كثيرة – لا يمكن للمرء أن يقول بأن الاختبار الصفي الكلاسيكي ليس الطريقة الصحيحة الوحيدة التي تطور التعلم الفعّال. فهناك حاجة إلى بحوث أكثر، مع التأكيد على أهداف وغايات الاختبارات، وعلى طرق لقياس ولمقارنة مستويات التعلم في الميدان الوجداني.

SUMMARY الخلاصة

إن المهمة الأولى التي تواجه الباحثين هو اختيار وصياغة مشكلة ما. والمشكلة البحثية عبارة عن سؤال أو بيان حول العلاقة بين المتغيرات. وفي محاولتهم لإيجاد مشكلة بحثية، قابلة للبحث قد يرجع الباحثون إلى خبراتهم الذاتية، وإلى نظريات يمكن استنباط أسئلة منها، أو إلى الأدبيات الراهنة في بعض ميادين الاهتمام. ويتوجب على الباحثين تقييم أهمية المشكلة المقترحة حسب معايير محددة وطرح أسئلة مثل: (هل ستساهم المشكلة في المعرفة القائمة حالياً؟ هل لدى المشكلة إمكانية التوجيه نحو بحوث أخرى؟ هل يمكن اختبارها – وهذا معناه – هل يمكن ملاحظة وقياس المتغيرات؟ ما مدى ملائمة المشكلة بالنسبة إلى اهتماماتي، وخبرتي، ومعرفتي بميدان البحث؟ هل يمكنني الوصول إلى البيانات اللازمة للمشكلة، وهل الأدوات متوفرة، أو هل يمكن بناؤها لقياس المتغيرات؟ هل يمكن تحليل البيانات وتفسيرها خلال الوقت المتاح؟) ويجب على المسألة أن لا تكون متضمنة قضايا فلسفية، أو يجب ألا تكون عامة بدرجة تجعل القيام بالبحث مستحيلاً. ويجب على بيان المسألة أن يحدد المجتمع الإحصائي ذا الأهمية وكذلك المتغيرات المزمع بحثها.

Key Concepts مفاهيم أساسية

criteria for research problem statement	معايير صياغة المشكلة البحثية
dependent variable	متغير تابع
independent variable	متغير مستقل
population	المجتمع الإحصائي
research problem	مشكلة بحثية
sources of research problems	مصادر المشكلات البحثية

EXERCISES تمارين

1. جد تقرير بحث منشور في إحدى المجلات الدورية، ثم أجب على الأسئلة التالية بالاستناد إلى قراءتك:
   أ. ما المشكلة التي يجري استقصاؤها في الدراسة؟
   ب. ما الفرضية (أو الفرضيات)؟
   ج. ما هي المتغيرات المستقلة والمتغيرات التابعة؟
   د. أين عثرت على المشكلة والفرضية (أو الفرضيات) الموجودة في التقرير؟

هـ. هل كانت المشكلة والفرضية (أو الفرضيات) معروضة بوضوح كاف بما يجعلك تدرك تماماً ما كان يجري استقصاؤه في الدراسة؟

2. اختر حقلاً واسعاً يمكن أن تهتم به في قيامك ببحث ما. ثم اختر إحدى أوجه هذا الحقل الواسع، وحدد مشكلة بحثية تثير اهتمامك في تتبعها. ثم أكتب المشكلة بصيغة معدة للبحث.

3. الأمثلة التالية بيانات غير ملائمة لمشكلات بحثية. أعد صياغة كل منها بحيث تصبح سؤالاً محدداً وملائماً للبحث.

أ. التدريب والأداء في اختبار الاستعداد الدراسي (SAT).

ب. المفهوم الذاتي الأكاديمي في تعليم اليافعين المعوقين.

ج. توقع التحصيل في السنة الأولى في كلية الحقوق.

د. تأثير الطريقة (x) في تدريس المفاهيم الرياضية.

هـ. اختلافات الجنس وقدرات التفكير الرياضي للمتميزين من ذوي أعمار ما قبل المراهقة.

4. قوّم المشكلات البحثية التالية:

أ. هل للفلسفة المتسامحة في تربية الطفل للدكتور "Spock" تأثير مضاد على التربية الأمريكية؟

ب. ما هي العلاقة بين الطريقة المفضلة لعقد الساقين عند الجلوس وبين ذكاء طالبات الجامعة الأمريكيات؟

ج. إذا تمعنا في الدراسات التجريبية الحديثة للتحصيل في الرياضيات، فهل يتوجب استبعاد "الرياضيات الحديثة" في المدرسة الابتدائية واستبدالها بالرياضيات التقليدية؟

د. كيف يتصور التلاميذ دور المدير في المدرسة المتوسطة المركزية؟

5. أذكر المتغيرات المستقلة والتابعة الأكثر احتمالاً في الدراسات الآتية:

أ. تأثيرات معالجة النصوص في الكتابة الشاملة والمراجعة لدى تلامذة الصف السادس؟

ب. حجز أطفال الروضات: النتائج السلوكية والأكاديمية في خلال نهاية الصف الثاني؟

ج. تأثيرات تدريس مجموعتين أو الصف بكامله على التحصيل في الرياضيات بعد إعادة تجميع التلاميذ في المرحلة الابتدائية.

د. التنبؤ بالمقررات الدراسية الأكاديمية في المدرسة الثانوية: الدور الذي يلعبه حجم المدرسة.

6. دون مصادر المشكلات البحثية التي استخدمها باحثون، وأعط مثالاً لكل منها.

7. ما تأثير قراءة الأطفال للكتب الجيدة على مهاراتهم القرائية؟ هل هذا السؤال قابل للبحث. وإذا كانت إجابتك بالنفي، كيف تقترح تغييره ليصبح صالحاً للبحث تجريبياً؟

الأجوبة ANSWERS

1. إن الأجوبة ستتنوع.

2. إن الإجابة ستتنوع.

3. أ. ما تأثير التدريب المحدد مقابل عدم التدريب للطلبة الذين يتهيئون لاختبار الاستعداد الدراسي، على أدائهم في الاختبار؟
ب. كيف يمكن للمرء وصف المفهوم الأكاديمي للذات لمجموعة من اليافعين المعوقين؟
ج. ما هي العلاقة لمتغيرات معينة محددة (مثل معدل النقاط) والدرجات في السنة الأولى في كلية الحقوق؟
د. ما هي العلاقة بين نوع الجنس وبين التفكير الرياضي للمتميزين من ذوي أعمار ما قبل سن المراهقة؟

4. أ. إن هذا السؤال يتضمن حكماً قيمياً يستحيل بحثه تجريبياً.
ب. هذا السؤال تافه، والإجابة عليه تسهم قليلاً في ميدان المعرفة.
ج. لا يمكن للبحث الإجابة على أسئلة قيمية، ويمكنه فقط تقديم معلومات تستند إليها القرارات.
د. بالرغم من إمكانية استقصاء المسألة، إلا أنها محدودة جداً، ولا تسمح في تعميمها إلى مواقف أخرى.

5. أ. متغير مستقل: معالجة النصوص (اللغوية) مقابل أسلوب بديل.
متغير تابع: قياس للقدرة الكتابية، أو تغير في القدرة الكتابية.
ب. متغير مستقل: تم حجزه أو لم يحجز في روضة الأطفال
متغير تابع: قياسات للتحصيل الأكاديمي والمميزات السلوكية في نهاية السنة الثانية.
ج. متغير مستقل: طريقة التجميع المستخدمة.
متغير تابع: التحصيل في الرياضيات.
د. متغير مستقل: حجم المدرسة الثانوية.
متغير تابع: عروض المقررات الأكاديمية في المدرسة (العدد، والتنوع).

6. الخبرة اليومية للمربين.
الاستنتاجات من النظرية.
الأدبيات ذات الصلة.
مصادر غير تربوية.

7. المصطلحات (كتب جيدة) و (مهارات القراءة) تحتاج إلى تعريف. يمكن إعادة كتابتها كالآتي:
ما تأثير قراءة الأطفال لأربعة كتب كلاسيكية مختارة على نمو المفردات اللغوية لديهم؟

REFERENCES المصادر

Burton, R.V. (1963). Generality of honesty reconsidered. Psychological Review, 70, 481-499.

Dewev, J. (1933). How we think. Boston: Heath.

Gamoran, A. (1993). Is ability grouping equitable? Education Digest, 58, 44-46.

George, P.S. (1993). Tracking and ability grouping in the middle school: Ten tentative truths-Middle School Journal, 24, 17-24.

Hartshorne, H., and May, M.A. (1928). Studies in the nature of character: Studies in deceit. New York: Macmillan.

Hunt, J.McV (1965). Traditional personality theory in the light of recent evidence. American Scientist, 53, 80-96.

Jussim, L. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual biases, and accuracy. Journal of Personality and Social Psychology, 57, 469-480.

Kourilsky, M., and Wittrock, M.C. (1992). Generative teaching: An enhancement strategy for the learning of economics in cooperative groups. American Educational Research Journal, 29, 861-876.

Marsh, R. (1984). A comparison of take-home versus in-class exams. Journal of Educational Research, 78, 111-113.

Nelsen, E.A., Grinder, R.E., and Mutterer, M.L. (1969). Sources of variance in behavioral measures of honesty in temptation situations: Methodological analysis. Developmental Psychology, 1, 265-279.

Pallas, A.M., Natriello, G., and McDill, E.L. (1989). The changing nature of the disadvantaged population: Current dimensions and future trends. Educational Researcher, 18(5), 16-22.

Piaget, J. (1932). The moral judgment of the child. Glencoe, IL: Free Press.

Popham, J. (1991). A slice of advice. Educational Researcher, 20(9), 18, 35.

Rosenthal, R., and Jacobson, L. (1968). Pygmalion in the classroom. New York: Holt, Rinehart and Winston.

Schmidt, P. (1993). Debate over ability grouping gains high profile. Education Week, 12, 23.

U.S. Department of Labor, Secretary's Commission on Achieving Necessary Skills. (1991). What work requires of schools. Washington, DC.

Wittrock, M.C. (1974). Learning as a generative process. Educational Psychologist, 11, 87-95.

الفصل الثالث

مراجعة الأدبيات
Reviewing The Literature

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون بوسع الطالب أن:

1- يكتب قائمة الوظائف الرئيسة المستخدمة في البحوث من خلال مطالعة الأدبيات.

2- يصف المصادر المرجعية الرئيسة في التربية.

3- يصف مؤشرات (ERIC)، ويوضح اختلافاتها واوجه التشابه فيما بينها.

4- يخطط طريقة عامة لاستعمال مؤشرات دورية في البحث.

5- يسمي ما لا يقل عن ثلاثة مؤشرات خارج (ERIC)، ويعطي أمثلة على نوع المعلومات الموجودة في كل منها:

6- يكتب الكيفية التي يستخدم بها نظام (SSCI) في مطالعة الأدبيات، ويفسر عمل كل واحدة من المؤشرات الأربعة.

7- يصف استخدام (الكتب السنوية للقياسات العقلية) كمصدر للمعلومات المتعلقة بالاختبارات وطرق القياس (أي التي تستعمل في المقياس).

8- يستخدم الحاسوب لتقصي الأدبيات ذات العلاقة في موضوع محدد أو خاص.

9- يقدم سلسلة من الخطوات التفصيلية المنهجية لتنظيم الأدبيات، ويفسر الهدف من كل خطوة.

وإذا ما تم تحديد أو تعريف إحدى المشكلات، عندئذ يكون الباحث متلهفاً بصورة طبيعية للعمل. وبالرغم من كل شيء، فإن من الخطأ الانغماس رأساً في تخطيط وتنفيذ البحث في واحد من ميادين الاهتمام. فالعنوان يجب أن يكون على صلة بالمعرفة الملائمة في ذلك الحقل. ومن الأهمية بمكان بالنسبة للمربين وللآخرين المنشغلين بالبحث أن يعرفوا كيف يجدون، وينظمون، ويستخدمون الأدبيات في ميادينهم. ويناقش هذا الفصل (1) دور الأدبيات ذات الصلة في أحد المشاريع البحثية. (2) مصادر المراجع في التربية، و (3) مهمة تنظيم الأدبيات ذات الصلة لتقديمها في التقرير.

دور الأدبيات ذات الصلة في مشروع بحثي
THE ROLE OF RELATED LITERATURE IN A RESEARCH PROJECT

يجب أن يتم البحث عن الأدبيات ذات الصلة قبل المباشرة الفعلية بالدراسة. وهذه المرحلة تفيد عدة وظائف مهمة:

1- المعرفة بالبحث ذي الصلة تمكن الباحثين من تقرير حدود مجال بحثهم. وباستخدام التشابه قد يقول أحد المكتشفين ((إنا نعلم بأن وراء هذا النهر سهولاً تمتد (2000) ميل غرباً، وخلف تلك السهول هناك سلسلة جبال، لكننا لا نعلم ماذا يقع وراء تلك الجبال. أنني اقترح عبور تلك السهول، وصعود الجبال، ومن ثم الاستمرار في السير غرباً)) وبالطريقة نفسها قد يقول الباحث: ((إن أعمال (أ، ب، جـ) قد كشفت هذا المقدار حول سؤالي، وإن بحوث (د) قد أضافت مقدارا آخر إلى معرفتنا، وأنني اقترح أن أذهب إلى ما هو أبعد من عمل (د) بالطريقة التالية)).

2- إن مراجعة دقيقة للنظرية وللبحوث ذات الصلة يمكّن الباحثين من وضع أسئلتهم وفقاً للوضع الصحيح. ويجب على المرء أن يقرر فيما إذا كان من المحتمل لمحاولاته أن تضيف إلى المعرفة بطريقة مفيدة. فالمعرفة في أي مجال معلوم تتكون من النتائج المتراكمة للعديد من الدراسات التي قامت بها أجيال من الباحثين، ومن النظريات المرسومة لدمج هذه المعارف ولشرح هذه الظواهر الملاحظة. ويتحتم على المرء مراجعة الأدبيات بهدف إيجاد الصلة بين دراسته وبين المعرفة المتراكمة في ميدان الاهتمام لذلك المرء. أما الدراسات التي لا رابطة لها بالمعرفة القائمة، فنادراً ما تقدم إسهامات مهمة إلى الميدان. فدراسات كهذه تميل لإنتاج أجزاء معزولة من معلومات ذات فائدة محدودة.

3- إن مراجعة الأدبيات ذات الصلة يساعد الباحثين على تحديد سؤالهم، وتوضيح

وتعريف مفاهيم الدراسة. ويمكن إن يكون السؤال البحثي واسعاً أكثر مما ينبغي مما لا يساعد الباحث على الاستمرار، أو أن يكون غامضاً أكثر مما ينبغي بحيث لا يمكن العمل فيه. ومراجعة دقيقة للأدبيات تقدم العون للباحثين في تنقيح أسئلتهم الأولية بحيث يمكن بحثها. كما إن المراجعة تساعد أيضاً في توضيح المفاهيم البنائية المستخدمة في الدراسة، وفي ترجمة تلك المفاهيم البنائية إلى تعاريف إجرائية. والكثير من هذه المفاهيم البنائية التربوية والسلوكية – مثل الإجهاد، والإبداع، والإحباط، والروح العدائية، والتحصيل، والدافعية، والتكيف – تحتاج إلى توضيح وإلى تعريفها إجرائياً. فهذه، بالإضافة إلى العديد من المفاهيم البنائية التربوية والسلوكية – لا تساعد كثيراً في البحث العلمي ما لم يجر تحديدها كمياً لقياسها. وتتيح مراجعة الأدبيات، معرفة المرء بالجهود السابقة لتوضيح هذه المفاهيم البنائية وتعريها إجرائياً. وعمليا ما ينتج عن المراجعة الناجحة تكوين فرضيات بخصوص العلاقات بين متغيرات الدراسة: ويمكن لهذه الفرضيات أن تقدم الإرشاد والتركيز للدراسة.

4- إن مراجعة نقدية للأدبيات ذات الصلة، غالباً ما تؤدي إلى نفاذ البصيرة في فهم أسباب النتائج المتناقضة في أحد ميادين البحث. فالنتائج المتناقضة أمر مألوف. والأسباب الكامنة وراء عدم الثبات قد يعود إلى أنواع الأساليب المتبعة في إيجاد حلول المشكلة، أو في أنواع الأدوات المستخدمة، أو في طرق البحث المستخدمة، أو في التحليلات. وقد توضح المقارنة بين أساليب هذه الدراسات عدم ثبات النتائج. وحل مثل هذه التناقضات بشكل تحديا، لكن ذلك سوف يكون مساهمة مهمة في المعرفة الخاصة بحقل اهتمام الباحث.

فمثلاً، وأثناء بحوث سابقة، وجدت (Bardwell, 1984) ما يؤيد الاقتناع بأن الفشل يضعف التعلم، وأخرى تؤيد وجهة النظر المعاكسة القائلة بأن الفشل يسهل التعلم. وافترضت بأن الفشل سيكون أكثر إضعافاً لأولئك الذين يتصورون بأنهم قد فشلوا في الدراسة أكثر من أولئك الذين لا يتصورون أنهم فشلوا. لذلك قامت بتصميم دراسة لاختبار هذه الفرضية. وقد أعطي طلبة الكلية أولاً ثلاثين مشكلة تحتاج إلى محاكمة عقلية لا يمكن حلها. ثم سئلوا فيما إذا كانوا يظنون بأنهم نجحوا أم أخفقوا. بعد ذلك أعطي لأولئك الأفراد ثلاثين سؤالاً أخرى يمكن حلها. وخلافا لفرضيتها فقد نال أولئك الذين ظنوا بأنهم أخفقوا في المشكلات المستعصية درجات أعلى في المشكلات اللاحقة القابلة للحل، مقارنة مع أولئك الذين قالوا بأنهم قد نجحوا. وقد كان التفسير المؤقت (غير النهائي) للباحثة لتلك النتائج يقول بأن الإخفاق سهّل التعلم لأولئك الذين ظنوا بأنهم أخفقوا في المشكلات مستحيلة الحل وذلك باستثارتهم نحو محاولة جادة في المشكلات اللاحقة القابلة للحل، بينما أولئك الذين ظنوا بأنهم قد نجحوا في المشكلات مستحيلة الحل لم تجر استثارتهم نحو محاولة جادة في حل المشكلات القابلة للحل.

5- من خلال دراسة البحوث ذات الصلة، يدرك الباحثون طرق البحث المفيدة، وتلك التي تبدو أقل فائدة. وسيقوم الباحث بتطوير سعة إطلاعه خلال غوصه في البحوث التي

تعرضها الأدبيات ذات الصلة. وعندما يتعمق المرء في الموضوع، يبدو واضحاً أن نوعية البحوث تتغير على نحو بالغ. وفي نهاية الأمر، يتحتم على الباحث البدء بملاحظة أن جميع الدراسات في أي ميدان ليست متكافئة بالضرورة. ويصبح المرء، بعد قليل، ناقداً للدراسات، وملاحظاً للطرق التي يمكن بها تحسين تلك الدراسات. فمثلاً، غالبا ما تبدو الدراسات الأولى في أي حقل معين غير متقنة وعقيمة. ولهذا يجرى، على الدوام، تحسين منهج البحث والتصميم مع كل دراسة جديدة. وبالرغم من ذلك، فالكثير من مشاريع البحث غالباً ما تفشل بسبب استخدام أساليب أو أدوات أو تصميمات أو تحليلات إحصائية غير مناسبة. وسيكتشف الباحث المسار الصحيح لاتباعه، من خلال اختباره المتعمق للدراسات الجيدة وإدراك الرديئة منها.

6- الاستقصاء الشامل للبحوث ذات الصلة، يجنّب التكرار غير المقصود لدراسات سابقة. فغالباً ما يقوم أحد الباحثين بتطوير فكرة قيمة، إلا إنه يكتشف أن دراسة مماثلة جداً قد جرى بحثها سابقاً. وفي هذه الحالة يتوجب على الباحث أن يقرر فيما إذا كان سيتعمد إعادة العمل السابق أو تبديل الخطط المقترحة والبحث في وجه مختلف من أوجه المشكلة.

7- إن دراسة الأدبيات ذات الصلة يضع الباحثين في موقع أفضل لتفسير أهمية نتائج أعمالهم. الإلمام بنظرية في حقل، إضافة للبحوث السابقة، تتيح للباحث إحكام نتائج بحوثه لتصبح جزءاً من المعرفة في ذلك الحقل.

مصادر المراجع في التربية
REFERENCE SOURCES IN EDUCATION

لاشك، فإن الأمر الجوهري للعلماء والباحثين يكون في معرفة كيف يعثرون على أعمال سابقة في حقولهم. ومن أجل القيام بذلك، يتوجب على المرء أن يعرف: (1) مصادر الأعمال السابقة، (2) وكالات جمع معلومات كهذه وتنظيمها في قاعدة معلومات "database"، (3) الشكل الذي تكون عليها قواعد البيانات، (4) الطرق الكفيلة بإيجاد المعلومات التي يحتاجها المرء. ومن أجل استخدام هذه المصادر، يتوجب على المرء أن يكون ملماً بالتسهيلات المتوفرة لدى المكتبة وخدماتها. ونموذجيا، توجد لدى المكتبات أدلة مطبوعة تصف خدماتها وأنظمتها وتجدول جولات إرشادية. وتستخدم الكثير منها فهرسا بالحاسوب (On line) متصلاً بنظام حاسوب تلك المكتبة. ومهم جداً على وجه الخصوص تعلم كيفية تنظيم فهرس البطاقات أو الحاسوب. وكقاعدة عامة، فإن المكتبات تقدم تعليمات تصف استخدام المبتكرات التكنولوجية المتنوعة وقواعد المعلومات وغيرها، باستخدام الحاسوب أو الأقراص المدمجة. ويتحتم على المرء أن يجد فيما إذا كان بوسع المكتبة الحصول على الكتب والمواد الأخرى – مثل الأطروحات والمقالات الدورية – من مؤسسات أخرى وذلك باستخدام ((نظام الإعارة الداخلية بين المكتبات)).

وأغلب المكتبات العامة الآن لديها ((الخدمة السريعة للوثائق)) من خلال نظام الإعارة الداخلي الذي يسمح للمكتبة بالحصول على مقالات الدوريات خلال يوم أو اثنين عن طريق وكالات خاصة بتسليم الوثائق.

أدلة أساسية Basic Guides

ومن أجل البدء في البحث عن دراسات بحثية، فمن المفيد العودة إلى أدلة أساسية مختصة بالأدبيات البحثية. إن أحد هذه الأدلة المفيدة يدعى ((دليل وصفي لقائمة المراجع للبحث التربوي (Berry, 1990)). والقصد من هذا أن يكون (كدليل مختصر لمساعدة الطالب في المقررات التربوية بالقيام باستخدام مفيد للمراجع في كليته أو في جامعته) وهو عبارة عن قوائم بأسماء أكثر من (700) مصدر بحثي مرتبة حسب نوع المواد، مثل الدوريات، والدراسات البحثية، والمنشورات الحكومية، والمواد المرجعية، والاختبارات. وتشتمل المواد المرجعية على الكتب السنوية، والأدلة، والمصادر الوصفية لقائمة المراجع وكتيبات عن منهج البحوث التربوية، وعلى شكل وأسلوب كتابة الأوراق البحثية.

أما الدليل العائد لمنظمة اليونسكو والمسمى (المصادر الحالية للمراجع التربوية / 1984)، فقد كان بثلاث لغات (الإنكليزية/ الفرنسية/ الإسبانية)، وهو يركز على مجموعة مراجع في المكتب الدولي لمركز الوثاق التعليمية لمنظمة اليونسكو في باريس. وهو يتضمن بيلوغرافيا راهنة (قائمة مراجع) وفهارس للدوريات، وسجلات بحث وهي مرتبة حسب البلدان، مع قوائم منفصلة لمنشورات تعود لمنظمات محلية وعالمية.

أما ((دليل مصادر المعلومات التربوية (Woodbury, 1982))) فهو دليل شامل لإيجاد المعلومات في التربية. فهذا الكتاب يختار ويصف، وفي حالات كثيرة يقيم الطبعة الرئيسة، وغير المطبوعة، والمصادر المؤسساتية للبحوث التربوية. ويستند الترتيب أساساً على نوع المطبوعة، إضافة إلى فصول عن المال والحكومة، والتربية الخاصة، والمواد التعليمية، والاختبارات وأدوات التقييم. كما يتضمن توضيحاً للخطوات المتخذة في البحث المفيد، ودليلاً عن الباحثين.

أما ((دليل إلى الكتب المرجعية/ 1986)) لمؤلفه (Sheehy)، و ((دليل إلى الكتب المرجعية: الذي يغطي المواد من 1985-1990)) لمؤلفه (Balay, 1992)، فهما يصفان باختصار يقيمان عدة آلاف من المراجع التي تخص حقولاً لموضوعات متنوعة تشمل التربية. ويجري تحديث عمل (Balay) مرة كل سنتين. أما ((الكتاب السنوي للمراجع الأمريكية)) لمؤلفه (Waynar) والمنشور منذ 1970 فإنه يغطي كتباً مرجعية (وطبعات معادة) في جميع الموضوعات لكل السنوات السابقة، كما يقدم ملحوظات وصفية وتقييمية.

مراجعات للأدبيات ذات الصلة بالتربية

Reviews of Education-Related Literature

المراجعات النقدية التي تقوم بتلخيص ودمج دراسات بحثية تدور حول موضوعات معينة يمكن أن تساعد في دفع أي شخص للابتداء في استقصاء الأدبيات. وهنالك عدة مراجعات أساسية يمكن للمرء العودة إليها. فموسوعة البحث التربوي (Alkin, 1992) المصممة لتقديم ((تركيب نقدي وتفسير للبحوث التربوية المنشورة)) تحتوي على مقالات تحمل تواقيع مؤلفيها إضافة إلى ثبت بالمراجع التي تقدم نقاشات موثقة توثيقاً جيداً للاتجاهات الحالية والتطورات وكذلك لموضوعات تقليدية. وفي هذه الموسوعة هناك ما يقارب (200) موضوعا تتضمنها هذه الموسوعة ذات الأجزاء الأربعة. وهذه الموسوعة مصدر أساسي جيد للإطلاع الابتدائي على البحوث في الميادين المتنوعة. أما **الموسوعة العالمية للتقييم التربوي** (Walberg & Haertel, 1990) فقد أصبحت واحدة من الأعمال المميزة في حقل التقييم التربوي. وهذه الموسوعة مقسمة إلى ثمانية فصول واسعة تغطي الدراسات التقييمية، وتقييم المناهج ونظرية القياس، وتطبيقات القياس ونماذج من الاختبارات والامتحانات ومنهج البحث، والسياسية التربوية والتخطيط. وهذا عمل أولي في إبراز نقاش العلماء حول مفاهيم أساسية في جميع مجالات التقييم التربوي. أما **موسوعة المربين الأمريكيين** (Dejnozka & Kapel, 1991) فإنها تضم حوالي (2000) من المدخلات القصيرة للمصطلحات والأسماء التي غالباً ما توجد في أدبيات التعليم المهني. **والموسوعة العالمية للتربية: البحث والدراسات** (Husen & Postlethwaite, 1994) فإنها تقدم: "مسحاً حديثاً للمعرفة الدقيقة المتصلة بالمشكلات التربوية، والممارسات، والمعاهد في كل أنحاء المعمورة". وهي تغطي عملاً ثقافياً ومهنياً في التربية، وتقدم مسحا للتطورات الحالية في الفروع المتنوعة للتربية، وتوفر معلومات صحيحة وسليمة علميا فيما يختص بتلك التطورات، ونماذج لبحوث أخرى نحتاجها. أما **موسوعة التعليم العالي العالمي** (Altabch, 1991) فهي مقتطفات مختارة رائعة في ميدان التعليم العالي العالمي، وتضم (67) مقالة حول قضايا كبيرة، وأفكار لموضوعات وأمم ومناطق جغرافية. وهناك **الموسوعة العالمية للتعليم العالي** (Knowles, 1978)، وهي مجموعة شاملة تتألف من عشرة مجلدات تضم (282) مقالة عن موضوعات معاصرة في التعليم العالي كتبها قادة تربويون من جميع مناطق العالم. أما **موسوعة التعليم العالي** (klark & Neave, 1992) فهي مصدر مماثل آخر، وتحتوي على أكثر من (300) مقالة قام بكتابتها علماء من أقطار كثيرة، وهي تسعى نحو تكامل بعيد الأثر "للمعرفة العالمية الراهنة حول التعليم العالي".

وهناك مصدر مراجعة مفيد آخر أسمه: **دليل البحث في التعليم** (Gage, 1963; Travers, 1973; Wittrock, 1985). وقد ظهرت ثلاث طبعات مختلفة، بين كل طبعة والتي تليها عشر سنوات. وهي تدون، وتلخص، وتحلل البحوث نقدياً في حقل التعليم. وكل نسخة منها تحتوي على مقالات جديرة بالثقة قدمها أناس اختصاصيون في موضوعات مختارة في الميدان. وهناك بيبلوغرافيا

شاملة لموضوعات مختارة أيضاً. ومن بين الموضوعات الموجودة في الطبعة الثالثة الآتي: (قياس التعليم) و (طرق بحثية كمية في التعليم) و (طرق نوعية للبحث في التعليم) و (الملاحظة بكونها استقصاء وطريقة) و (ودمج البحوث حول التعليم) و (تعليم استراتيجيات التعليم) و (سلوك المعلم وتحصيل الطالب) و (بحث في تعليم الفنون وعلم الجمال).

أما موسوعة التعليم الخاص (Mann, 1987) فإنها تقدم معلومات أساسية عن العديد من المهن وفروع الدراسة الخاصة بتعليم أطفال استثنائيين(*) وميزاتهم الذاتية وحاجاتهم ومشكلاتهم. وهذه الموسوعة المكونة من ثلاثة أجزاء تحتوي على أكثر من (2000) موضوعاً. ومع أن ترتيب هذا الكتاب هو أبجدي، إلا أنه يمكن تجميع الموضوعات المتنوعة في حقول سبعة، هي: السير الذاتية، اختبارات نفسية وتربوية، التوسط وتسليم الخدمات وحالات الإعاقة وخدمات ذات صلة، وقضايا قانونية، ومتنوعات أخرى. أما "التقدم في التربية الخاصة" (Keogh) التي يتم نشرها سنوياً منذ 1980، فتقدم اختبارات شاملة في حقول كثيرة وخلاصة تقييمات محددة في مجالات خاصة في التربية الخاصة.

وهنالك نشرة دورية فصلية لرابطة البحث التربوي الأمريكي (AERA) تدعى: **مراجعة البحوث التربوية**. فمنذ سنة 1931، كانت ولا تزال تنشر: (مراجعات تكاملية وتفسيرية لأدبيات البحث التربوي في قضايا منهجية أساسية). وفي محاولة لتقديم ملخصات للبحوث في مجالات واسعة للموضوعات في المهنة، بدأت رابطة البحث التربوي الأمريكي بنشر سلسلة (**مراجعة للبحوث في التربية**) سنوياً منذ عام 1973. وتسعى هذه السلسلة لتحديد ما أنجزته البحوث، وما هو قيد الإنجاز، وتلبية حاجات الحقل. وعلى سبيل المثال، فإن الجزء العشرين (1994) مرتب حول ثلاثة موضوعات رئيسة: تعليم المعرفة والممارسة، قضايا العدالة في النمو التربوي والتقييم، والمنهجية. كما إنها تتضمن مراجعات لحقول التعلم، والتعليم وعمل المعلمين، والعدالة والهوية، والتقييم والبحث.

إن الدراسات النفسية تلعب دوراً مهماً في البحث التربوي. والنشرة المسماة: (المراجعة السنوية لعلم النفس) التي تصدر سنوياً منذ 1950 تستخدم أخصائيين في المواد الدراسية لنشر تقييم أدبيات البحوث، والاتجاهات والتطورات الحديثة في جميع أوجه علم النفس كل عام، كما تشير أيضاً إلى المناطق المهملة. أما الكتاب المسمى: (دليل علم النفس العام) لمؤلفه (Wolman, 1973) فهو مراجعة أخرى مفيدة، ومصدر يتعامل مع المجالات الواسعة في التاريخ، والنظريات، والمنهجية، والكائن الإنساني، والإدراك والتعلم، واللغة والتفكير، والذكاء، والدافعية والعاطفة والشخصية ومجالات خاصة أخرى. كما تتضمن أيضاً مراجع ببيلوغرافية للدراسات البحثية الموجودة.

(*) طفل استثنائي (exceptional Child): طفل يختلف عن مستوى الطفل العادي من ناحية جسمية أو عقلية أو انفعالية أو اجتماعية إلى درجة تستلتزم معاملة خاصة به "د. محمد على الخولي". (المراجع)

أدلة/فهارس دورية، ومجالات دورية للخلاصات، وفهارس الشواهد
Periodical Indexes, Abstract Journals, And Citation Indexes

بعد القيام بتأسيس قاعدة واسعة للبحوث ذات الصلة، والنظريات والآراء حول موضوع الاهتمام، يمكن للمرء أن يبدأ بتحديد مواقع مواد إضافية لا يُستشهد بها في مصادر المراجعة الأساسية. فالأدلة الدورية المتنوعة، والمجلات الدورية للخلاصات والأدلة الاستشهادية (الشواهد) تمكن المرء من العثور على هذه المعلومات. إن هذه المنشورات التي تظهر في أجزاء متعاقبة وعلى فترات منتظمة تساعد كدليل لإيجاد المعلومات المنتشرة بصورة متناثرة في المجلات الدورية والمصادر الأخرى. إن ناشري هذه الفهارس/الأدلة يستخدمون قارئين محترفين حيث يقوم هؤلاء بمسح وتصنيف البحوث في مصادر منشورة وغير منشورة. وبعد ذلك يجري تجميع تلك البحوث حسب الموضوع وبعض الأحيان بالحواشي، وذلك من أجل تقديم قوائم شاملة وحديثة للباحثين بالأعمال في ميادين عملهم.

أدلة/فهارس إرك The ERIC Indexes

هنالك اثنان من أهم الفهارس الموجودة وينتجهما "مركز معلومات المصادر التربوية" – إرك – (ERIC) الذي أنشأته وزارة التعليم الأمريكية (USOE) من اجل جمع، وخزن ونشر المعلومات عن التربية [يشير (الخولي: قاموس التربية) إلى ERIC بوصفه: مركز المعلومات عن الموارد التربوية – المراجع].

وقبل ظهور نظام (ERIC)، فإن التقارير المقدمة إلى وزارة التربية الأمريكية من المتعهدين ومتلقي المنح كانت توزع أولياً بصورة متناثرة وبعدها تختفي كما جرى لتقارير من مصادر أخرى. وقد كان المقصود من نظام (ERIC) تصحيح هذا الموقف الفوضوي وجمع هذه المواد غير المنشورة و (الهائمة) ذات الأهمية للمربين، ومن أجل جعل هذا الذخيرة من المعلومات في متناول الناس.

إن نظام (ERIC) الذي يموله المعهد القومي للتربية (NIE)، يقوم بتجميع، وتقييم، وفهرسة، وتلخيص المعلومات لتضمينها في فهارس (ERIC) من خلال عملية تسهيل مركزية وشبكة من ستة عشر من دور المعاوضة(*) الموجودة في جامعات متنوعة ومنظمات مهنية في جميع أنحاء البلاد. وكل دار معاوضة مسؤولة عن منطقة تربوية معينة. ويسجل **دليل خدمات المعلومات ERIC** (Fustukjian & Taheri, 1992) هذه الدور وكل منظمة تزود مستفيديها وزبائنها بحرية الوصول إلى قواعد بيانات ERIC ومصادرها ذات الصلة.

(*) دار المعاوضة (Clearing house): مؤسسة تتلقى المعلومات والبيانات من مصادر مختلفة وتزود تلك المصادر بما لديها من معلومات على سبيل المعاوضة. "د. محمد علي الخولي: قاموس التربية، 1985" – (المراجع)

إن أهم ثلاث دوريات على الإطلاق التي يقوم بنشرها مركز ERIC هي: **الفهرس الراهن للمجلات الدورية في مجال التربية (CIJE)، ومصادر في التربية (RIE)، ومصادر تربية الأطفال الاستثنائيين (ECER)**. ويستخدم نظام فهرسة تم تطويره بعناية من اجل الوصول إلى وثائق ERIC. وهذا النظام قد جرى وصفه في إحدى النشرات المسماة "موسوعة ERIC للمترادفات" أو "Thesaurus Of Eric Descriptors" (Houston, 1995) وهي موجودة في أغلب المكتبات العامة.

الفهرست الراهن للمجلات الدورية في التربية

Current Index to Journals In Education

إن المجلة الشهرية (الفهرست الراهن للمجلات الدورية في التربية) (CIJE) يجري تجميعها من أعمال المتخصصين في دور المعاوضة التابعة لمركز (ERIC). وقد بدأت منشورات هذا الفهرس في 1969، ويجري تراكم محتوياته سنوياً أو نصف سنوي. وطبقاً للنظام الذي طورته مترادفات (ERIC) يجري تجميع المقالات من (800) مجلة دورية وتصنف وتفهرس. وهناك حالياً أكثر من (450000) مقالة قد جرى فهرستها في (CIJE) وفي كل سنة تضاف 18000 مقالة تقريباً.

والفهرس (EIJE) ينقسم إلى أربعة أقسام: فهرس الموضوعات؛ فهرس المؤلفين، قسم المدخلات الأساسية، وفهرس محتويات المجلات الدورية. ويمكن للمرء أن يجد مقالات يهتم بها عندما يبحث في فهرس الموضوعات عن العناوين وعن إعداد المقالات ذات الصلة، وبعد ذلك يستخدم تلك الأعداد لإيجاد المدخلة في قسم المدخلات الأساسي. إن مفردات موسوعة (ERIC) للمترادفات قد جرى تحديدها بإمعان لعرض نظام فهرسة أكثر نظاماً، ونتيجة لذلك، قد لا يجد المرء الجدول لموضوع خاص. وفي تلك الحالة، يتوجب على المرء أن يبحث عن مرادفات لذلك الموضوع الخاص. أما فهرس المؤلفين المنفصل فهو مفيد في إيجاد عمل أحد الباحثين على وجه الخصوص. ويشير فهرس محتويات المجلات الدورية إلى المجلات المشمولة وإلى محتويات كل إصدار منها. وتقدم (ERIC) خوارزمية(*) في ست خطوات توضح استخدام (UIJE).

المصادر في ميدان التربية Resources In Education

إن خلاصات التقارير البحثية الآتية من مصادر أخرى غير المجلات الدورية تجري فهرستها وتنشرها شهرياً (ERIC) في نشرة تدعى مصادر في التربية (RIE). كما يجري نشر طبعة تراكمية نصف سنوية من أقسام الفهرست أيضاً. وتقوم كل دار معاوضة تابعة لنظام (ERIC)

(*) خوارزمية (Algorithm): مصطلح يشير إلى مجموعات من الخطوات المتسلسلة التي ينبغي إتباعها، وهي مشتقة من عالم الرياضيات "الخوارزمي". (المراجع)

بتجميع المواد ذات الصلة بميدانها الاختصاصي ثم تفهرس الوثائق وتعد الأدلة والملخصات. أما الوثائق الأصلية فإنها ترسل، إضافة إلى الملخصات التي قامت بتحضيرها كل دار معاوضة، على أساس منتظم إلى مركز تسهيلات المعالجة المركزية لتصبح كلها جزءاً من قاعدة البيانات المركزية ليستفيد منها كل من يستخدم نظام (ERIC). إن مركز تسهيلات المعالجة المركزية مسؤول عن جمع كل ما يرد إليه من دور المعاوضة، كما ينشر الملخصات المتجمعة لديه شهرياً في "مصادر في التربية: RIE".

إن المصادر الموجودة لدى دور المعاوضة المتخصصة، تتضمن تقارير من جميع البحوث التربوية التي تمولها الحكومة الفدرالية، إضافة إلى الملخصات، والكتيبات، وأدلة المنهج، وأوراق متميزة من المعاهد والجمعيات العلمية، وبيليوغرافيا ومواد مقرر دراسي نموذجي، وأدلة مساعدة للمعلمين، وملخصات برامج، بالإضافة إلى خطط بحثية متنوعة وتقارير عن مشروعات بحثية. وبمثل هذه التغطية الواسعة، يتضح أن الوثائق الموجودة في مجموعات (ERIC) ذات أهمية حيوية للممارسين وللباحثين في التربية. وهناك أكثر من (350,000) وثيقة، والمجموعة تتزايد بمعدل (13,000) وثيقة في كل سنة.

ويتعين لكل ملخص الرقم الذي يعود له بهدف التحديد ولخدمات طلبات استنساخ الوثائق الأصلية. وهنالك ثلاثة فهارس: المؤلف، المؤسسة، الموضوع. وعند استخدام الفهرس المناسب، يجد المرء عناوين الوثائق وأرقام الوصول إليها حيث يستخدمها المرء للعثور على الملخصات التي يبحث عنها.

والإجراء المستخدم في مصادر التربية (RIC) يشابه ذلك المستخدم في الفهرست الحالي للمجلات الدورية في التربية (CIJE). فالمرء، نموذجياً، يبتدئ بأحدث نسخة ويعود بعدئذ إلى الخلف باحثاً عن الموضوعات ذات الاهتمام.

المصادر التربوية للطفل الاستثنائي Exceptional Child Education Resources

المجلة – التي كانت سابقاً تصدر تحت أسم (**الملخصات التربوية للطفل الاستثنائي**) – هي الآن نشرة ربع فصلية باسم "المصادر التربوية للطفل الاستثنائي: ECER" وتحتوي على ملخصات مخزونة في ملف الحاسوب العائد إلى مجلس مركز المعلومات للأطفال الاستثنائيين كجزء من برنامج (ERIC). وتجري فهرسة الملخصات العائدة لتربية الأطفال المعوقين والموهوبين والشباب. أما ترتيبها فهو مشابه لذلك الموجود في (CUE) و (RIE). وقد تراكمت الفهارس داخل كل مجلد خلال 1975، وابتداء من 1976، يمكن العثور على فهارس المؤلفين، والعناوين، والموضوعات في الإصدار الأخير من السنة. أما تغطية (ECER) فيشمل الكتب، ومقالات مجالات دورية ووثائق حكومية، وغيرها.

مجموعة وثائق (ERIC) The ERIC Document Collection

يمكن الحصول على نسخ ميكروفيش(*) للوثائق الأصلية إذا دعت الحاجة إليها، أو يمكن الطلب من (ERIC) استلام جميع الوثائق المطلوبة حالما تتوفر. وكثير من المكتبات تمتلك ترتيباً كهذا مع (ERIC) وبذلك تكون المجموعة الكاملة للوثائق متاحة للاستخدام في وقت مناسب.

وأغلب الوثائق متاحة في شكلين: النسخة المطبوعة والميكروفيش. وفي أغلب المكتبات فإن الميكروفيش هي المفضلة لأنها أقل كلفة وتحتاج إلى مساحة أقل لخزنها. والمجموعة الكاملة لوثائق (ERIC) على الميكروفيش تحتل مساحة عدد قليل من خزانات الأضبارات الصغيرة. وهناك جهاز يسمى (قارئ الميكروفيش) يستخدم لتكبير صور الصفحات على شاشة مندمجة به، كما يسمح الجهاز للقارئ بأن ينتقل من صورة صفحة إلى أخرى. وأغلب المكتبات والجامعات تمتلك مثل تلك الأجهزة. إن كلفة كل وثيقة من مجموعة (ERIC) موجودة في (RIE) لكل من الميكروفيش والنسخة المطبوعة.

إن جميع المشروعات التي تمولها وزارة التربية الأمريكية (USDE) موجودة بصورة أوتوماتيكية في نظام (ERIC) والنسخ الكاملة للتقارير عن هذه المشروعات متاحة من خلال (ERIC).

وعندما يتم النشر الأولي بموجب امتيازات حقوق النشر تكون هناك إشارة إلى المصدر، ومع ذلك، فإن النسخ لا تكون عادة متوفرة من خلال ERIC. ويصح الأمر ذاته بالنسبة للكتب والمواد الأخرى المعدة للبيع. إن (ERIC) لا يقدر بثمن بالنسبة للتربية، إذ أنه يجمع المعلومات ثم يلخصها من مصادر واسعة المجال ومن مواد وثيقة الصلة بالبحوث التربوية، التي كان أغلبها في السابق صعب المنال.

أما دور المعاوضة التابعة لمركز المعلومات أعلاه فتقوم بإعداد ببليوغرافيا مصحوبة بحاشية تفسيرية لموضوعات مختارة. وقبل أن يُعد المرء الببليوغرافيا الخاصة به، فمن الأفضل له أن يكتب إلى دار المعاوضة المختصة بحقل الاهتمام ليرى فيما إذا كان لدى العاملين ببليوغرافيا مفيدة لذلك المرء.

ويمكن إيجاد وصف أكثر شمولية لنظام (ERIC) بالإضافة إلى كيفية استخدامه في كتيب يدعى (ERIC: ماذا يمكن أن يقدمه لك، وكيفية استخدامه)، (Brown, Setts, & Yarborough, 1977) ويتوفر في مكتبات عامة كثيرة. ومع مرور عشرين سنة على نشره تقريباً، إلا أنه المصدر المتاح الوحيد.

(*) ميكروفيش: شريحة صغيرة تحمل صورا مصغرة لما بين 60-100 صفحة من كتاب "د. محمد علي الخولي: قاموس التربية" - (المراجع)

وترغب دور المعاوضة التابعة لشبكة (ERIC) بأن تجعل مجموعتها شاملة بقدر الإمكان، ولذلك فهي تطالب التربويين بأن يقدموا لها المواد المناسبة. وإذا ما قام المرء بإعداد تقرير، أو خطاب، أو ورقة بحث يرغب مؤلفها في نشرها عن طريق (ERIC)، فعليه أن يرسل نسخة إلى أحد فروع (ERIC).

كيفية الوصول إلى ERIC بواسطة الحاسوب

Computer Access to the ERIC System

قبل جيل من الآن كان البحث الكامل عن المجلات الدولية والوثائق المناسبة لمواد تتعلق بمشكلة بحثية معطاة، مهمة صعبة. أما اليوم، فإن الوصول بواسطة الحاسوب إلى نظام (ERIC) يجعل من المستطاع القيام ببحوث كهذه بكفاءة، باستخدام حاسوب شخصي أو بالاتصال الهاتفي.

إن محتويات كلي من (CIJE) و (RIE) متاحة على أشرطة حاسوبية، ويمكن القيام ببحث حاسوبي باستخدام الأشرطة. فمثلاً، لتحديد مقالات المجلات ووثائق (RIE) التي تهتم باستخدام مواد مبرمجة لتعلم اللغة الفرنسية، يمكن للحاسوب تحديد جميع الوثائق والمقالات باستخدام الأمر الواصف (descriptor): (التعليم المبرمج) والأمر (الفرنسية). وبعض برامج الحاسوب تستخدم أرقام الوصول الخاصة بكل من (ED & EJ)، لكن أغلب البرامج تستخدم مدخلات CIJE الكاملة، وكذلك مدخلات RIE الكاملة لكل مقالة ووثيقة لهما أمر واصف.

وتمتلك أكثر من (900) مؤسسة أشرطة (ERIC)، وبالإمكان استخدام الحاسوب (online) للبحث. ولدى أغلب دوائر التربية الحكومية الأشرطة أعلاه، ويتيسر للتربويين الإفادة من خدمات الحاسوب دون مقابل. ولدى أغلب المؤسسات اختصاصيون في استرجاع المعلومات وفي استطاعتهم مساعدة الزبائن في ترجمة المطالب العامة حسب الشكل اللازم للحاسوب لاسترجاع المعلومات المطلوبة من نظام (ERIC) والكثير من الجامعات تقدم خدمات مشابهة إلى طلبتها والمستفيدين الآخرين.

ولدى بعض المكتبات ERIC وقواعد بيانات(*) على أقراص مدمجو (CD-ROM) كطريقة حديثة للوصول إلى قواعد بيانات ERIC أو غيرها. وقد قامت مؤسسة (Silver Platter Information) بعرض قواعد البينات الكاملة لنظام ERIC الممتدة منذ سنة 1966 حتى الآن على قرصين مدمجين (CD-ROM). ويحتوي كل قرص على استشهادات بيبلوغرافية، وملخصات، وأوامر واصفة. وهذه الأقراص يمكن حملها داخل حقيبة كتب صغيرة، خلافاً لفهارس ERIC المطبوعة. ويمكن استخدام الأقراص عن طريق حاسوب شخصي. وبالإضافة إلى استخدام قاعدة بيانات ERIC بالحاسوب، هنالك حوالي (800) قاعدة بيانات أخرى تضطلع بموضوعات أخرى متنوعة، وبذلك يشمل (ملخصات سايكولوجية) وملخصات لأطروحات

(*) قاعدة بيانات (database): مصطلح يشير إلى مجموعة البيانات التي يخزنها الحاسوب حيث يمكن استخدامها بيسر، كما يمكن الإضافة إليها (المراجع).

عالمية، وملخصات سوسيولوجية، وفهرست للأستشهادات في علم الاجتماع. أما قاموس Gale لقواعد المعلومات (Marcaccio, 1994) فإنه يقدم أوصافاً كاملة للإنتاجات الإلكترونية لقواعد البيانات، كما يزود معلومات اتصال بالبائعين والموزعين، وشروط الاستخدام، وقوائم مفردات عروضهم.

وهناك قواعد بيانات حديثة يتم إنتاجها بمعدل شبه يومي لنظام البحث الحاسوبي ويتوفر لدى منافذ البيع. و "دليل فهرس التربية" واحد من العديد من قواعد البيانات التي تشتمل عليها خدمة الحاسوب الخاصة بنظام (Wilson line). وتتوفر هذه الخدمة في العديد من المكتبات.

ويعد التنقيب في نظام ERIC خطوة مهمة لاقتفاء الأدبيات ذات الصلة، غير أنه لا يمكن للباحث الافتراض بأن هذه الخطوة كافية للبحث حال الانتهاء منها، إذ ربما لا تكون المادة المناسبة للسؤال قد دخلت نظام ERIC.

فهارس الدوريات الأخرى Other Periodical Indexes

هنالك الكثير من فهارس الدوريات في حقل التربية، وهي مفيدة لتحديد معلومات جديدة تتصل بالبحث، إضافة إلى الأفكار المعاصرة. وكما ذكرنا سابقاً، فإن أحد الفهارس المعيارية/ المقننة للحقل هو (فهرست التربية)، الذي تنشره شركة (H. W, Wilson) منذ عام 1929. وهذا الفهرست يعرض مقالات من (400) دورية، والكتب السنوية، والنشرات والوقائع(*)، وسلاسل المقالات المختصة بموضوع واحد. وتسعة وثمانون دورية من أصل (400) دورية المذكورة ليست مسجلة في (CIJE). ويعد (فهرس التربية) أفضل المصادر لتحديد مواقع مقالات المجلات الدورية التي تم نشرها قبل إنشاء CIJE (الفهرس الراهن للمجلات الدورية في مجال التربية) في 1969، وكذلك لكل للمقالات الحديثة، إذ أنه يدون بصورة منتظمة مقالات قبل أن يقوم (CIJE) بتدوينها بحوالي ستة أشهر. والقصور الملاحظ في (فهرس التربية) هو أنه لا يحتوى أية هوامش.

وبالإضافة إلى هذه الفهارس العامة، هنالك عدد من الفهارس المتخصصة متيسرة لاستخدام الباحث. وفهرس (ملخصات سيكولوجية) يقوم بتدوين كل ما يعرض في العالم من أدبيات علم النفس وبقية فروع المعرفة ذات الصلة. وهذا الفهرس نشر لأول مرة في 1927، وهو يتضمن كتباً، وأطروحات الدكتوراه، ومقالات دورية، وخلاصة لكل منها، وبذلك يتمكن القارئ من تقرير مدى صلته بالمادة. أما فهرس (ملخصات تطور الطفل، وبيليوغرافيا) الذي ظهر أيضاً في 1927 فإنه يقدم مؤلفاً وموضوعاً في مجالات الطفولة، والطب السريري والصحة العامة، والإرشاد، وعلم النفس التطوري والمقارن، والتجريبي. أما فهرس (ملخصات

(*) الوقائع Proceedings: مصطلح يشير إلى خلاصات التقارير أو السجلات المنشورة للمناقشات أو اللقاءات أو المؤتمرات – (المراجع)

في علم الاجتماع التربوي) فهو نشرة عالمية، يدون المقالات الدورية حسب الموضوع والمؤلف، والكتب والأطروحات في ميدانه، مع خلاصة لكل منها. ويزود فهرس (ملخصات الإدارة التربوية) المجلات المتخصصة في ميدان الإدارة التربوية بدراسات، مستخدماً أسلوب الموضوع والمؤلف. أما "ملخصات التعليم العالي" فهي تجميع لملخصات من المجلات الدورية، والمؤتمرات والوقائع، وتقارير البحوث ذات الصلة بطلبة الكليات وخدمات الطلبة. والموضوعات تشمل الإرشاد والسكن، والمعونات المالية والاختبارات والقياس. ويدون (فهرس العلوم الاجتماعية)، بطريقة المؤلف والموضوع، مقالات المجلات الدورية في حقل العلوم الاجتماعية، ويشمل المجلات التربوية، وأخرى من حقول علم النفس، وعلم الاجتماع، وعلم السياسة، وعلم الأجناس، والقانون، وعلم الاقتصاد.

وتوجد هناك فهارس متخصصة لكل فرع من فروع المعرفة، يتضمن التربية الأعمالية(*)، والفنون الصناعية، والطب، وقد ذكرنا القليل منها. والاستعانة بالأدلة الأساسية في الأدبيات سوف يعطي القارئ أسماء الفهارس المتخصصة في الحقول الأخرى التي قد يحتاجها.

فهارس رسائل الماجستير وأطروحات الدكتوراه

Indexes of Dissertations and Theses

تعد رسائل الماجستير وأطروحات الدكتوراه مصادر مفيدة للمعلومات بالنسبة للباحثين. وبالإضافة إلى فهارس الدورية ومجلات الملخصات التي تدون بعض الرسائل والأطروحات، مثل RIE في ERIC وملخصات علم نفسية والعديد من أدلة متخصصة وفهارس أيضاً مفيدة للغاية.

و(**الفهرس الشامل للأطروحات**) (1861-1972) يحاول شمول جميع الأطروحات في الولايات المتحدة من 1861 وحتى 1972. وهذا الفهرست يحتوي على سبعة وثلاثون مجلداً في سبعة عشر تصنيفاً، إضافة إلى استشهادات كاملة بيبلوغرافية في تدوين المؤلف والموضوع. وهناك تجديد سنوي ملحق لهذا الفهرست.

وفهرست (**ملخصات أطروحات عالمية**) الذي بدأ نشره عام 1938- يحتوي على ملخصات لأطروحات دكتوراه كانت قد قدمته إلى (الشركة العالمية للأفلام المصغرة الجامعية) جامعات متعاونة (حوالي (550) في عام 1993). إن (ملخصات أطروحات عالمية) مقسم إلى ثلاثة أجزاء: الإنسانيات والعلوم الاجتماعية، العلوم والهندسة، والجامعات الأوروبية. ويمكن البحث بواسطة الحاسوب في (ملخصات أطروحات عالمية) من خلال عدة منافذ كبرى للبيع عن طريق الحاسوب. وأغلب المكتبات العامة تقدم تسهيلات للباحثين باستخدام قاعدة بيانات حاسوبية.

(*) التربية الأعمالية Business Education: تربية تهدف إلى تعريف الطالب بمبادئ إدارة الأعمال وفهم نظام العمل والاقتصاد والتجارة والعلاقات العامة (د. محمد علي الخولي: قاموس التربية) – (المراجع)

وهناك نشرة أخرى مشابهة، اسمها: (**أطروحات دكتوراه أمريكية**). وهي تبلور في قائمة واحدة أطروحات أعطيت درجاتها العلمية في أمريكا وكندا خلال سنة أكاديمية ما، وكذلك تلك الأطروحات الموجودة على الأفلام الصغيرة (مايكروفيلم). كما تتضمن هذه النشرة أيضاً عدداً من الأطروحات غير مذكورة في (ملخصات أطروحات عالمية). ونشرة (أطروحات دكتوراه أمريكية) مرتبة حسب الموضوع والمؤسسة، ولديها فهرس مؤلفين، لكن ليس لديها ملخصات.

أما نشرة (**ملخصات رسائل جامعية: ملخصات لرسائل ماجستير منتقاة ومطبوعة على أفلام صغيرة / 1986**) فإنها تحتوي على ملخصات موجودة على مايكروفيلم لقائمة منتقاة من مقالات الرسائل من جامعات وكليات أمريكية متنوعة. ومن أجل تدوين رسائل جامعية صادرة عن كليات وجامعات أمريكية، متخصصة في حقل التربية، فهنالك مصدر (Silvey): (رسائل ماجستير في التربية). وهو يصدر سنوياً منذ عام 1951، ولا يوجد فيه ملخصات، لكنه يقدم معلومات بيبلوغرافية عن كل رسالة مذكورة لديه.

فهرس الشواهد في ميدان العلوم الاجتماعية Social Sciences Citation Index

عندما يعثر المرء على مقالات مهمة وذات صلة، فيمكنه استعمال فهرس الشواهد من أجل التقدم إلى أمام في حينه. فمثلاً، إذا ما قرأ شخص ما مقالة مفيدة منشورة في 1986، فيتمكن ذلك الشخص من خلال فهارس لاحقة من تحديد مقالات أحدث من التي تستشهد بتلك المقالة خلال البيبلوغرافيا العائدة لتلك الفهارس. والكثير من تلك المقالات سوف تقدم وصفاً للاستمرار أو ا لامتداد أو التعديل في المقالة الأصلية ذات الاهتمام.

إن "**فهرس الشواهد للعلوم الاجتماعية**" (SSCI) ذي المجلدات الثلاثة سنوياً والذي ظهر في 1973 في (معهد المعلومات العلمية) يقوم بتحديد المؤلفين الذي تم الاستشهاد بأعمالهم خلال السنة في جميع ميادين العلوم الاجتماعية، وبضمنها التربية، وما قد كتب في حقول متنوعة. كما يشتمل على معلومات بيبلوغرافية ضرورية لكل من المؤلفين الذين كان قد استشهد بأعمالهم أو القائمين بعملية ذكر الشواهد نفسها. وتتوفر هذه المعلومات بواسطة أربعة فهارس هي :

1- (فهرس المصادر) وهو عبارة عن قائمة أبجدية بجميع المؤلفين المنشورة أسماؤهم في المجلات التي يشتمل عليها (فهرس الشواهد للعلوم الاجتماعية) خلال السنة. وهنالك معلومات بيبلوغرافية كاملة عن المقالات المنشورة لأولئك المؤلفين، يتبعها قائمة أبجدية بالمؤلفين الأوائل لكل من المصادر المستشهد بها في كل مقالة. ويشير (فهرس المصادر) إلى المؤلفين الثانويين أيضاً.

2- (فهرس الشواهد) وهذا يقدم قائمة أبجدية بالمؤلفين الأوائل لجميع الأعمال المستشهد بها في

المقالات المذكورة في (فهرس المصادر)، ويعقبها السنة، والمجلة، والمجلد، والصفحة الأولى من كل مقالة كتبها ذلك المؤلف. ولكل مقالة، هناك أسماء الأشخاص الذين قاموا بالأستشهاد بها، ويتبعها المجلة، والمجلد، والصفحة الأولى، والسنة التي حدث فيها ذلك الاستشهاد. وهكذا يسمح هذا الفهرست لأي شخص بتتبع العمل في ميدان معين، أو بتتبع مؤلف محدد.

3- (فهرس تبادل مصطلحات الموضوع) يأخذ كل كلمة مهمة ويقرنها مع كل كلمة مهمة أخرى في كل عنوان. وكل كلمة في عنوان ما يجري بعدئذ تدوينها كمصطلح أولي يرتبط بكل من المصطلحات الأخرى بوصفها مصطلحات مرافقة. وهناك قائمة أبجدية بأسماء المؤلفين الذين تحتوي عناوين مقالاتهم على الكلمات حيث يتم ربطها بالمصطلح الأولى والمصطلح المرافق له. ويمكن، بعدئذ، إيجاد معلومات بيبلوغرافية لكل مؤلف في (فهرس المصادر).

4- (فهرس العنوان المتحد) هو قائمة أبجدية للمنظمات التي ينشر فيها المؤلفون نتاجاتهم خلال السنة. وتحت كل مدخلة متحدة توجد قائمة بأسماء المؤلفين مزودة بمعلومات بيبلوغرافية كاملة.

المنشورات الحكومية Government Publications

الحكومة الفيدرالية – باعتبارها مصدراً رئيسياً للمعلومات التربوية – ترعى بحوثاً أكثر، وتقوم بمسوحات أكثر، وتجمع إحصائيات لجميع الأنواع أكثر من أية منظمة أخرى في الولايات المتحدة. إن وزارة التربية – وكذلك المركز القومي للإحصائيات التربوية – يقوم بتوزيع أعداد هائلة من المنشورات وذلك يتضمن التقارير البحثية، والمسوحات، والأعمال الإدارية، والبرامج الوصفية.

ولإيجاد منشورات حكومية أمريكية معينة، فإن (البيان الشهري للمنشورات الحكومية الأمريكية) هو أفضل فهرس يمكن الركون إليه. والقسم الأساسي فيه يدون الوثائق التي تنشرها كل وكالة حكومية، ويتضمن العنوان، والمؤلف، والموضوع، وفهارس الكلمات الدليلية. وقد كانت هناك تجميعات سنوية مرتبة بالعنوان والموضوع تصدر في شهر كانون الأول/ ديسمبر حتى عام 1975، لكن منذ عام 1976 أصبح هناك تجميع كل ستة شهور. أما منشورات الوزارات الحكومية الأخرى للتربية، والوكالات الحكومية الأخرى، فيمكن تحديد مواقعها من خلال (القائمة الشهرية للمنشورات الحكومية).

مصادر الاختبارات Test Sources

نحتاج الاختبار أو وسيلة القياس، غالباً، لدى القيام بالبحث. و (معهد بورروس Buros

للقياسات العقلية) بنشر (الكتب السنوية للقياس العقلي): وهي مصادر مرجعية مهمة تدون وتراجع الاختبارات نقدياً. وقد بدأ نشرها منذ 1938، وهذه الكتب مخصصة لمساعدة مستخدميها في ميدان التربية، وعلم النفس، والصناعة، وذلك من اجل استخدام أبرع للاختبارات المقننة. وكل كتاب سنوي مرتب على نفس النمط، والقصد منها تكملة – وليس نسخ – المجلدات السابقة. والاختبارات مرتبة حسب الموضوع، ومواصفات أي اختبار يتبعها مراجعات نقدية، وإشارات إلى دراسات استخدمت فيها تلك الاختبارات. ويحتوى كل مجلد إحالات إلى مراجعات، واقتباسات، وإشارات بيبلوغرافية في مجلدات سابقة. وتشمل المجلدات اختبارات استعداد وتحصيل في ميادين متنوعة، واختبارات الشخصية، واختبارات مهنية، واختبارات الذكاء، بالإضافة إلى معلومات كاملة لكل اختبار، وذلك يتضمن الكلفة المادية وتعليمات طلبها. وتعتبر "اختبارات معروضة للبيع 17" (Murphy, Conoley, & Imparta, 1994) كفهرس وملحق للكتب السنوية الأحد عشر الأولى في القياسات العقلية. ويقدم (معهد بورس) أيضاً خدمة قاعدة بيانات حاسوبية، مع تجديد المعلومات شهرياً. والمعالم الكاملة تسمح للمستفيدين بالبحث وطبع النص كاملاً لكل وثيقة، أو تصفح النصوص وطبع أجزاء مختارة من السجلات المسترجعة. وهذا نظام استرجاع ممتاز للمعلومات، حيث التجديد في المعلومات مستمر حتى بين فترات ظهور كتاب السنوية.

ويعد كتاب (Sweetland & Keyser, 1986) بعنوان "اختبارات: مرجع شامل للتقويم في علم النفس والتربية والأعمال" إضافة إلى ملحقه، مرجعاً متميزاً، فهو يشتمل على (3000) اختبار، متاحاً باللغة الإنكليزية. وقد أعطي كل اختبار تصنيفاً أولياً، ووصفاً تفصيلياً في أحد الأجزاء، وقد تتم الإحالة إلى صنف آخر.

ومنذ سنة 1984، فإن كتاب (keyser & Sweetland) المسمى (مقالات نقدية في الاختبار) قدم مواصفات عامة ومقالات نقدية حول أدوات القياس، كما اشتمل على فهرس موضوع تراكمي للاختبارات حسب نمط المتغيرات. وهناك إرشادات متخصصة للاختبارات تشتمل على "مقاييس الاتجاهات النفسية الاجتماعية" (Robinson & Sharer, 1973)، و"قراءة في الاختبارات ومراجعات" (Buros, 1968-1975) و"دليل إلى 75 اختبار تشخيصي في ميدان التربية الخاصة" (Compton, 1984)، و "دليل عالمي للاختبارات المكانية(*) (Eliot & Macfarlane, 1983)"، و "تقييم المتعلم العاجز: أدوات مختارة" (Mauser, 1981)، و "اختبارات في المايكروفيش: فهرست ذو هوامش"، و "فهرس مجموعة اختبارات لخدمة الاختبارات التربوية".

(*) اختبار مكاني (Spatial Test): اختبار لقياس إدراك الطفل العلاقات المكانية. (د. محمد على الخولي: قاموس التربية" – (المراجع)

ويمكن إيجاد المعلومات حول الاختبارات في فهارس الدوريات، ويشمل ذلك (الفهرس الراهن للمجلات في التربية: CUE) و (فهرست التربية)، و (مصادر التربية للأطفال الاستثنائيين ECER)، (وملخصات سايكولوجية)، حيث تدون الاختبارات تحت اسمائها المحددة في فهرس الموضوعات. ويمكن للفرد أن يبحث تحت عناوين الموضوعات مثل (اختبارات) و (استعراض الاختبارات). كما يغطي (دليل المستهلك إلى الاختبارات المطبوعة: Ha-mmill, 1992) وهذا يسد ثغرة في أدبيات نقد الاختبارات إذ أنه المصدر الوحيد الذي تخضع فيه جميع الاختبارات الموجودة داخله إلى مراجعة باستخدام تقييم مقنن معين. وقد يفيد هذا الكتاب في اختيار أحد الاختبارات، أو في تدريس عملية تقييم الاختبارات، ولتحسين نوعية الاختبارات.

البحث باستخدام الحاسوب Computer Searching

توفر أغلب المكتبات الأكاديمية العامة خدمات بحث حاسوبية وتسمى أحياناً البحث الحاسوبي عن قواعد بيانات، أو خدمات مرجعية حاسوبية. والكثير من تلك المكتبات تقدم جلسات تدريبية لتعليم المستفيدين كيفية القيام بالبحث بأنفسهم. ولدى أغلب المكتبات اختصاصيون يتمتعون بالمهارة في تقصي قواعد البيانات لخدمة الزبائن. وقد تحتاج بعض المكتبات إلى حضور الزبون عندما يكون اختصاصي المعلومات يقوم بالتقصي الحاسوبي.

وقد جرى دمج الكثير من المصادر المشار إليها في هذا الفصل في قواعد البيانات الحاسوبية، وذلك يتضمن (فهرست التربية) و (فهارس مركز معلومات المصادر التربوية :ERIC) و (المصادر التربوية للطفل الاستثنائي)، و (ملخصات سيكولوجية) و (فهرست شواهد العلوم الاجتماعية)، و (الملخصات العالمية لأطروحات الدكتوراه)، (الكتب السنوية للقياسات العقلية). وبالإضافة إلى ذلك، هنالك قواعد بيانات أخرى، مثل: (قاعدة بيانات مرض نقص المناعة المكتسب :AIDS)، و (قواعد بيانات بحوث العلوم الطبية)، و (مصادر في التدريب المهني :RIVE)، و (مجموعة اختبارات لخدمة الاختبارات التربوية :ETSF)، و (ملخصات سوسيولجية :SOCA). وللحصول على قوائم تفصيلية لمعظم قواعد البيانات المتاحة، يتحتم على المرء العودة إلى فهارس قاعدتي البيانات الاثنتين التي تنشرها (Dialog Information Services) و (تكنولوجيا المعلومات :BRS). وتستخدم الكثير من المعاهد أمناء مكتبة محترفين ماهرين في تقصي قواعد البيانات الحاسوبية لتقديم المساعدة في استخدام الأجهزة المختلفة.

يقدم البحث بواسطة الحاسوب عدة فوائد: أولها توفير الوقت. فعندما تكون هناك عدة مرادفات أو مصطلحات متشابهة للمفهوم الواحد، ففي دقائق يتمكن الحاسوب من استعادة وطبع المعلومات، بذلك يلغي الساعات أو حتى الأيام في حالة البحث اليدوي. ثانياً: أنه سائد. فقواعد البيانات تتجدد أسبوعياً أو مرتين في الأسبوع، ولذلك فهي أحدث من الفهارس المطبوعة، وثالثاً هي الارتباط. فعندما يكون موضوع البحث متضمناً موضوعين معاً أو أكثر

فيتمكن الحاسوب عندئذ من البحث في أكثر من موضوع واحد في الوقت ذاته. رابعاً: يتيح التقصي الحاسوبي دقة أكبر. فعندما لا يتيسر تحديد موضوع ما بسهولة في الفهارس المطبوعة، عندئذ يتيح الحاسوب البحث عن عناوين بديلة، أو رموز موضوعات أو كلمات رئيسية، أو رؤوس موضوعات إضافية لا تذكرها الفهارس التقليدية. وبوسع خدمة البحث الحاسوبي تجاوز المعلومات غير المهمة وذلك بتحديد عوامل مثل التاريخ ونوع المنشور، واسم المؤلف، أو اللغة. والفائدة الخامسة هي التفرد. فبعض قواعد البيانات تقدم معلومات لا تتوفر في المصادر المطبوعة. ووظائف الفرز لها فائدة مهمة أخرى. فيمكن لخدمة البحث الحاسوبي فرز المعلومات المسترجعة طبقاً إلى تاريخ نشرها (صعوداً أو نزولاً)، أو أبجدياً حسب المؤلف أو العنوان.

وبالنسبة لأغلب الباحثين فإن التقصي باستخدام الحاسوب هو الطريقة المختارة. فالتقصي اليدوي يوجب على المرء أن يتفحص عدة فهارس للدوريات من أجل عنوان معين، ثم يوجب عليه بأن يتتبع العنوان خلال الفهارس، إلى إن يعثر على مدخلات مناسبة وإلى دمج المصطلح المختار مع آخر ذي أهمية. وعلى أية حال، فالبحث بواسطة الحاسوب يتيح التنقيب عن عدة موضوعات في وقت واحد ثم يعود بتجميعها مستخدماً مفاهيم منطقية تعرف باسم (العوامل البولية، وذلك نسبة إلى العالم الرياضي الإنكليزي جورج بول "George Boole" الذي قام بتطوير نظام منطقي في القرن التاسع عشر عرف بعدئذ باسمه).

الشبكة الدولية للمعلومات The Internet

يعتبر الكثير من الباحثين (الشبكة العالمية للمعلومات) مصدراً مفيداً جداً. فالشبكة العالمية للمعلومات سلسلة من شبكات جامعية وحكومية ووطنية ودولية للمعلومات وجميعها موجودة في شبكة معلومات واحدة منطقية. وهي تمكن الباحثين من المشاركة في معلوماتهم وخبراتهم. والشبكة العالمية للمعلومات نظام معقد لنشر المعلومات وللبحث عن المعلومات، للاتصال بالآخرين، وللتعليم والتعلم، ومن أجل إدارة الأعمال التجارية. فهي تقدم بريداً إلكترونياً، وتحويلاً للملفات، وخدمات متنوعة أخرى. كما تيسر الوصول إلى الصحف الإلكترونية، والوثائق الحكومية المتنوعة، والمعلومات المفيدة لأجراء البحوث.

وعند ربط المعلومات، والمصادر الحاسوبية بالاستبصار الإنساني فيمكن للشبكة أن تساعد الباحثين في الوصول إلى قواعد متنوعة للبيانات الخاصة بموضوعات متعددة حول العالم بأجمعه. فمثلاً، للحصول على معلومات حول التقييم، يمكن للفرد أن يستعمل الشبكة للوصول إلى قواعد البيانات لدار المعارضة بمركز معلومات المصادر التربوية (ERIC) لهذا الغرض. والمصادر المتوفرة في دار المعاوضة هذه تشمل نظام الاختبارات التربوية (ETS) ومجموعة قواعد بيانات الاختبارات التي تقدم أحد أنظمة البحث عن أكثر من (9500) مواصفات اختبار، والكتب السنوية للقياسات العقلية ومواصفاتها وكذلك استعراض لاختبارات محددة.

القواميس Dictionaries

يمكن القول إن بوسع كل فرع من فروع المعرفة الأكاديمية، تقريباً، أن تكون له مفرداته اللغوية الخاصة به. وكل قاموس موضوع هو شرح الكلمات بطريقة موجزة - سواء كانت مصطلحات أو أسماء - وهي تكون اللغة الاصطلاحية لذلك الموضوع الخاص. وهناك عدة قواميس أساسية في حقل التربية. إن (قاموس التربية :Good, 1973) هو قاموس علمي للكلمات وللمصطلحات التي تملك معان خصوصية في حقل التربية. وهناك ما يقارب (200) من المنسقين، والمشاركين، والمراجعين الذين سطرت أسماؤهم في النسخة الثالثة من القاموس، التي تحتوي على أكثر من (33,000) مدخلة. والمصطلحات التربوية المستعملة في كندا وإنكلترا وويلز، قد جرى تعريفها في أجزاء منفصلة وذلك في نهاية القاموس. وبالرغم من أن القاموس (Good) يحمل تاريخ نشره إلا أنه يعتبر كلاسيكياً في حقله. وهناك (القاموس العالمي للتربية Page, Thomas, & Marshal, 1982) ويعتبر أكثر حداثة ، إذ يشمل أكثر من (10,000) مادة تغطي تعابير ومصطلحات، ومنظمات دولية، ومعاهد وطنية كبرى وجمعيات، ومربين، وما شابه ذلك.

أما (قاموس التربية :Hall, 1982) فيقدم مقالات وجيزة حول خمسة عشر حقلاً من حقول التربية، ويتبع ذلك قائمة أبجدية بالمصطلحات (وخاصة المصطلحات البريطانية). أما (الموسوعة الدولية للتعليم العالي :Husen & Postlethwaite, 1985) فإنها تشمل في جزئها الأول على تعاريف وجيزة للألفاظ الأوائلية(*)، ومعجم للمصطلحات. وفي حقل التربية الخاصة، هنالك (قاموس القراءة والعجز التعليمي :Bush & Andrews, 1978)، وهناك قاموس أسمه (مصدر المعلم ومرشده إلى التربية الخاصة: مصطلحات/ قوانين/ اختبارات/ منظمات :Davis, 1980)، وهناك (قاموس المصطلحات للتربية الخاصة Moore, Abraham & Lainge, 1980) وهذه القواميس الثلاثة أساسية وعملية.

مصادر إحصائية Statistical Sources

وبالنسبة للإحصائيات التربوية، فإن الحكومة الفدرالية الأمريكية، تتبعها الولايات والحكومات المحلية، تعد مسئولة عن أكبر عدد من الوثائق الإحصائية. أن كتاب (مصادر إحصائية O'Brien & Wasserman, 1994:) وهو كتاب إحصائي مرجعي- يعرض تقريباً (72,000) مصدراً إحصائياً تحت (20,000) من الموضوعات. وهو يمثل دليل موضوعات لبيانات حول عناوين متنوعة للولايات المتحدة والدول الأخرى. وطبعة 1993 من هذا الكتاب تحتوي أيضاً قوائم بأسماء قواعد بيانات إحصائية فيدرالية، وأرقام هواتف إحصاء فيدرالية. وهناك كتب مرجعية عديدة أخرى - مثل (التقويم العالمي وكتاب الحقائق) و (تقويم المعلومات) و (ملخصات

(*) اللفظية الأوائلية (Acronym): كلمة مركبة من أوائل حرف كلمات أخرى (المورد، 1992) - (المراجع)

إحصائية عن الولايات المتحدة الأمريكية) - وجميعها تحتوي على إحصائيات في حقول تربوية وفي موضوعات أخرى. وعلى العموم، فإن الإحصائيات موثوقة وهي مصادر لكثير من الإحصائيات الموجودة.

وهناك ثلاثة فهارس للإحصائيات قام بنشرها جهاز المعلومات التابع لمجلس النواب الأمريكي ومركزه في واشنطن، وهي (فهرس الإحصائيات الأمريكي :ASI) و (فهرس المراجع الأمريكية :SRI)، و (فهرس الإحصائيات الدولية :IIS). فالأول يفهرس ويلخص تقريباً كل مصدر إحصائي صادر عن الحكومة الفيدرالية الأمريكية، بينما يفهرس الثاني ويلخص الوثائق الصادرة عن الدولة. ويشتمل السابق على العديد من الإحصائيات غير الحكومية، وهي تلك الصادرة عن شركات وأعمال تجارية خاصة، وتلك الآتية من منظمات النفع العام ومن الاتحادات. أما الفهرس الثالث فيشتمل على إحصائيات حكومية كبرى من جميع أنحاء العالم. وهو مصدر ممتاز للأرقام الإحصائية في الأمم المتحدة. وهذه الفهارس الثلاثة موجودة أيضاً على أقراص ليزرية لدى جهاز المعلومات التابع لمجلس النواب الأمريكي. وبالإضافة إلى الفهارس الثلاثة المذكورة، هناك (فهرس التربية) وهو مصدر ثمين لإيجاد الإحصائيات التربية.

ويعد منذ سنة 1962 (خلاصة الإحصاء التربوي) مصدراً مهما - يغطي حقلاً واسعاً من التعليم الأمريكي من رياض الأطفال إلى كلية الدراسات العلياً. وهذا الكتاب يتضمن مختارات إحصائية من مصادر كثيرة - حكومية وخاصة. كما يعتمد على نحو خاص، على نتائج المسوح والنشاطات التي يقوم بها المركز القومي لإحصائيات التربية (NCES).

الاستعارة الداخلية بين المكتبات Interlibrary Loan

ويمكن لخدمات الاستعارة الداخلية بين المكتبات أن تكون مفيدة جداً. فإذا ما وجد شخص ما مصادر لدراسات بحثية في دوريات، أو كتب، أو في أطروحات غير متوفرة في المكتبة التي يتردد عليها الباحث، فبالوسع استعارة هذه المواد من مكتبات عامة أخرى من خلال خدمات الاستعارة الداخلية بين المكتبات. ويتوقع أن يضع المرء بالحسبان فترة أسبوعين اثنين عند التفكير في الاستفادة من خدمات الاستعارة الداخلية بين المكتبات.

تنظيم الأدبيات ذات العلاقة بالبحث
ORGANIZING THE RELATED LITERATURE

عندما يقتنع الباحث بأنه قام بدراسة شاملة ومعقولة للأدبيات في الحقل، فبوسعه متابعة مهمة تنظيمها. ويمكنه، كأسلوب مفيد، ترتيب الدراسات حسب الموضوع، وتقرير مدى صلة كل منها بدراسته الخاصة.

ويتوجب عليه أن يقاوم إغراء تقديم الأدبيات كسلسلة من الملخصات. بل يجب عليه وضع أساس نظامي للدراسة. فينبغي تقديم الأدبيات بطريقة تسوّغ قيامه بتلك الدراسة، وذلك بإبراز ما قد علم وما هو باق للبحث في الموضوع ذي الاهتمام. وتهيئ الفرضيات إطاراً لتنظيم كهذا. ومثله مثل المكتشف الذي يقترح حمله استكشافية فعليه وضع خطة توضح المنطقة المعروفة وتشير إلى طريق الوصول للمنطقة التي يود استكشافها. وإذا ما تضمنت الدراسة عدة أوجه، أو اتجهت لاختبار أكثر من فرضية واحدة، فإن عملية التنظيم تجري لكل واحدة منها بمعزل عن البقية.

ومن المحتم تقريباً، فإن عدداً من التقارير التي قام المرء بدراستها بعناية وشملها بملاحظته، سيتضح – عند الإمعان فيها – على أن صلتها بموضوعية هامشية. وليس ضروريا ولا مرغوباً فيه ذكر كل دراسة في مقترح البحث يكون الباحث قد واجهها خلال مطالعة الأدبيات. فالقارئ لا يخرج بانطباع جيد. فصلة المادة وتنظيمها هما من الأهمية بمكان.

إن الباحث الذي يفشل في تجميع الأدب ذي الصلة، بطريقة منظمة منذ البداية، قد يغدو مشوشاً، وقد تكون المقترحات التالية مفيدة:

1- ابدأ بأحدث الدراسات في مجال البحث، ثم تابع عملك رجوعاً إلى الخلف من خلال المؤلفات الأقدم. والفائدة الواضحة لهذا الأسلوب تكمن في أن المرء يبدأ بالدراسات التي دمجت أفكار ونتائج البحوث السابقة. فسوء الفهم السابق قد جرى تصحيحه، والأساليب غير المجدية قد جرى تجنبها. وهناك فائدة أخرى هي إن تلك الدراسات سوف تشمل مراجع لأعمال سابقة، ولذلك، فإنها سوف تقوده إلى مصادر قد لا يقع عليها في ظروف أخرى. ومن الواضح فإنه يجب أن يضع حدوداً لعملية جمع البحوث ذات الصلة. فمن الجانب الأول، إن وضع الأساس المفيد لأية دراسة يتطلب شمول جميع الأعمال المهمة في الحقل. ومن الجانب الثاني، فإن تكريس وقت مفرط لهذه المحاولة قد ينتج عنه ضجر لدى قارئ التقرير، بالتفصيلات غير الضرورية. وتأكد بأن الأدبيات ذات الصلة تخدم الموضوع لكنها لا تهيمن على عمل الفرد.

2- اقرأ الخلاصة أو فصول الملخص لأي تقرير ما، أولاً، لتقرر فيما إذا كان وثيق الصلة بالمشكلة. إن عمل كهذا قد يوفر وقتاً كثيرا بدلاً من إضاعته في قراءة مقالات غير ذات جدوى.

3- تصفح التقرير بسرعة قبل تدوين الملاحظات، لتحدد الفقرات التي لها صلة بالمشكلة – وهذه طريقة أخرى لتوفير وقت القراءة.

4- أكتب ملاحظاتك مباشرة على بطاقات فهرس، لأنها سهلة الترتيب والتنظيم من الأوراق الكبيرة. ويفضل الكثير من الباحثين بطاقات حجم 6×4 بوصة التي لها فراغات مناسبة لكتابة الملاحظات.

5- أكتب وصفاً بيبلوجرافياً كاملاً لكل عمل أدبي. وإذا أدرك المرء طبيعة نموذج الدليل المختصر الذي سيستخدم في التقرير النهائي فبوسعه توفير وقت باستخدام ذلك الشكل المرجعي لتدوين ملحوظاته. أضف إلى ذلك رقم الإيداع المكتبي حتى يسهل عليك إيجاد ذلك العمل الأدبي ثانية من الرف -- إن كان ذلك ضروريا.

6- تسهيلاً للتصنيف والتنظيم، لا تكتب أكثر من مرجع واحد على البطاقة الواحدة، فليس من السهل ترتيب المراجع أبجدياً، أو بأية طريقة أخرى، ما لم تكن مكتوبة بطريقة منفردة.

7- كن متأكداً بالإشارة إلى ما هو مقتبس من كلمات المؤلف مباشرة، وما الصياغة التي أعددها للمقتبسات. والإخفاق في تمييز ذلك قد يؤدي إلى انتحال غير متعمد. ومن الحكمة أيضا فصل تقييم المؤلف لبحثه عن النتائج التي توصل إليها الباحث.

8- إذا استخدم المرء قاعدة بيانات حاسوبية، ، فيجب الاحتفاظ باستراتيجيات البحث مكتوبة على الورق. فهذه المعلومات ستساعد في استعادة المعلومات وتقليل الكلفة المادية إن كان هناك ضرورة لتحديث المعلومات.

الخلاصة SUMMARY

إذا شمل الباحث سبل المعلومات في مجال بحثه كما اقترحنا، فلابد أن تكون النتيجة صورة كاملة معقولة للموقع الذي سوف تشغله الدراسة داخل مجال البحث . (وفي الصفحة التالية، تجد خلاصات عن معظم المصادر التربوية ومحتويات الأدبيات ذات الصلة في التربية، وسنعرضها في نصها الإنجليزي لمن يود الإفادة منها – المراجع).

مصادر مهمة للأدب ذي الصلة في حقل التربية

Important Sources Of Related Literature In Education

SOURCE	CONTENT
Child Development Abstracts and Bibliography	Abstracts of journal articles in the field of child development.
Current Index to Journals in Education	Citations of journal articles related to education, with annotations where needed. Indexed by subject, by author, and by journal. No abstracts.
Dissertation Abstracts International	Abstracts of doctoral dissertations in the United States, Canada, and Europe.
Education Index	Citations of journal articles related to

	education. Indexed by subject and title. No abstracts.
Encyclopedia of Educational Research	Summaries and evaluations of research published at the end of each decade.
ERIC Microfiche Collection	Complete documents, the abstracts of which are in Resources in Education.
Exceptional Child Education Resources	Abstracts of journal articles and other sources in special education.
Mental Measurement Yearbooks	Information and evaluations of commercially available tests.
Monthly Catalog of Government Publications	Subject listing of U.S. federal government publications.
Psychological Abstracts	Abstracts of journal articles in psychology.
Readers' Guide to Periodical Literature	Citations of articles in popular journals. Indexed by subject and title. No abstracts.
Resources in Education	Abstracts of research reports and other documents acquired by ERIC clearinghouses.
Review of Educational Research	Reviews on various topics in each quarterly issue.
Social Sciences Citation Index	Bibliographic information for cited authors and topics.
Test Critiques	Reviews of measurement instruments, including critiques, and technical aspects.
Tests in Print IV	Index and supplement to the first 11 Alental Measurement Yearbooks.
Thesaurus of ERIC Descriptors	Terms for classifying and indexing ERIC documents.

Key Concepts مفاهيم أساسية

Boolean logic	المنطق البولي
Computer searching	البحث باستخدام الحاسوب
ERIC system	نظام (مركز المعلومات والمصادر التعليمية)
Guides to governmental documents	أدلة الوثائق الحكومية
Indexes to dissertations and theses	فهارس أطروحات الدكتوراه والماجستير

فهرس الدوريات	Periodical index
قواميس متخصصة	Specialized dictionaries
مصادر إحصائية	Statistical sources
مصادر الاختبارات	Test sources

تمارين EXERCISES

1- أذكر ثلاثة أدوار مهمة للأدب ذي الصلة في مشروع بحثي.

2- قارن بين ميزات (الفهرس الراهن للمجلات الدورية في التربية CIJE) و (فهرس التربية).

3- أي من الفقرات التالية تعتبر إستراتيجيات جيدة لتنظيم بحث في الأدبيات؟

أ. نظّم الدراسات حسب العنوان.

ب. إبدأ بالمقالات المبكرة واستمر على العمل إلى الأمام في الوقت ذاته.

جـ. أقرأ الخلاصة لأي تقرير أولاً.

د. أقرأ بسرعة التقرير من أجل إيجاد الفقرات العائدة إلى المشكلة المختارة.

هـ. أكتب مرجعاً بيبلوجرافياً كاملاً لكل بحث.

4- من خلال مشروع البحث، في أي نقطة يتوجب على الباحث أن يقوم بمسح أدبي وبحث يخص السؤال المختار؟

5- ما معنى الحاشية.

6- وضح أهمية (الكتب السنوية للقياس العقلي) في تحديد مصادر الاختبار.

7- لماذا يعتبر (مركز معلومات المصادر التربوية ERIC) مصدراً مفيداً في البحث عن الأدبيات.

8- وضح استعمال (فهرس شواهد العلوم الاجتماعية) في بحث في الأدبيات.

9- ما هي الخطوات النموذجية في استخدام نظام ERIC في مراجعة الأدبيات.

10- ما هي مزايا استخدام الطريق الحاسوبي للوصول إلى قواعد بيانات كفوءة للأدبيات؟

الأجوبة ANSWERS

1- إن المعرفة بالبحث ذي الصلة تجعل الباحث قادراً على تعريف حدود الحقل، ووضع المشكلة في مكانها الصحيح، وتجنب التكرار غير المقصود لدراسات سابقة.

2- إن (فهرس التربية) يقدم وسيلة لتحديد مقالات نشرت من سنة 1929 إلى الوقت الحالي. أما (CIJE) فيغطي مجلات أكثر ويزود الحواشي، بيد أن فهرس التربية يدون المقالات أسرع من CIJE.

3- أ، جـ، د، و، هـ.

4- يقوم الباحث بمسح الأدبيات ذات الصلة حالما يتم اختيار مسألة البحث والتخطيط لها.

5- إن الحاشية تعني وصفاً وجيزاً لمقالة ما. وفي إمكانها مساعدة الباحث في اختيار مقالات حول موضوع بحثه.

6- إنها تحوي على أكثر القوائم الشاملة ووصفاً للاختبارات المقننة المتوفرة. أما المراجعات النقدية للاختبارات، فإنها تساعد المرء في اختيار اختبار مناسب.

7- إن (مركز معلومات المصادر التربوية ERIC) شمولي جداً في تغطيته لكثير من المجالات في التربية. وكل دار من دور المعاوضة تختص بأحد المجالات لأحد الموضوعات حيث يقوم بتجميع وإدخال المعلومات والمواد إلى داخل المنظومة. كما يقوم (ERIC) بتغطية واسعة لتشكيلة من المواد: مثل الدراسات البحثية، والمشروعات، وأدلة المناهج الدراسية، وأوراق مهمة من محاضر الجمعيات الثقافية والمعاهد، والبيبلوجرافيا، والمواد ذات الصلة بالمقررات الدراسية. إن الكثير من هذه المواد غير منشورة وغير موجودة في أي مكان آخر. إن (ERIC) لا يهيئ قوائم بأسماء المواد فقط، لكنه يقدمها مسجلة على ميكروفيتش.

8- من خلال (فهرس شواهد العلوم الاجتماعية) يتمكن المرء من تتبع مسار في البحث مبكر زمنيا ومن ثم تحديد مقالات قد اقتبست عملاً معيناً.

9- أ. قم بتحديد كلمات رئيسة في الدراسة.

ب. جد الكلمات الرئيسة أو مرادفاتها في موسوعة ERIC للمترادفات.

جـ. استخدام برنامجاً حاسوبياً لإيجاد أي من مقالات المجلات ووثائق ERIC، تحوي المصطلحات الرئيسة أو ما يصفها أو قم بالبحث اليدوي خلال CIJE و RIE لتجد الأرقام التي توصلك للمقالات والوثائق ذات المصطلحات الرئيسة. أبدأ بالنشرات الحديثة ثم أرجع إلى الخلف.

د. أقرأ الخلاصات والحواشي من اجل حذف المقالات والوثائق التي ليست لها أهمية.

هـ. جد المقالات المتبقية، إضافة إلى وثائق ERIC.

و. أقرأ أولاً الخلاصة، ثم الموجز، ثم الخاتمة.

10. يتمكن التفتيش الحاسوبي من البحث في نظم قواعد البيانات بسرعة. وهذه مفيدة على وجه الخصوص من أجل تحديد المقالات والمواد الأخرى التي لها كلمات رئيسة. وهي تمكن من الوصول بطرق متنوعة، منها بواسطة العنوان، أو التاريخ، أو المؤلفين أو الكلمات الرئيسة، أو رموز الموضوعات. وقواعد البيانات الحاسوبية أحدث من الأدلة المنشورة وتحوي معلومات لا تتوفر في أي مكان آخر.

REFERENCES المصادر

Abbott, T. (Ed.). (1994). Internet world's on Internet .94:An international guide to electronic journals, newsletters, texts, discussion lists, and other resources on the Internet. Westport, CT: MecKlermedia.

Alkin, M.C. (Ed.). (1992). Encyclopedia of educational research (6th ed., 4 vols.). New York: Macmillan.

Altbach, P.G. (Ed.). (1991). International higher education: An encyclopedia (2 vols.). Hamden, CT: Garland.

American Doctoral Dissertations. (1957-).Ann Arbor, MI: University Microfilms.

American Library Association. (1992). Library resources on the Internet: Strategies for selection and use. Chicago: American Library Association, Reference and Adult Services Division.

American Statistics Index. (1973-). Bethesda, MD: Congressional Information Service.

Annual Review of Psychology. (1950-).Stanford, CA: Annual Reviews.

Balay, R. (Ed.). (1992). Guide to reference books: Covering materials from 198.5-1990. Chicago: American Library Association.

Bardwell, R. (1984). Failure: Facilitating or debilitating? Journal of Experimental Education, 52, 192-194.

Berry, D.M. (1990). A bibliographic guide to educational research (3d ed.). Metuchen, NJ: Scarecrow Press.

Brown, J.W., Setts, M.K., and Yarborough, J. (1977). ERIC: What it can do for you, hou, to use it. Stanford, CA: ERIC Clearinghouse on Information Sources.

BRS Database Catalog. (1994). Latham, NY: BRS Information Technologies.

BRS Information Technologies, 8000 Westpark Drive, McLean, VA 22102.

Buros, O.K. (Ed.). (1938-).Mental measurements yearbook. Lincoln. NE: Buros Institute of Mental Measurements, University of Nebraska.

Buros, O.K. (Ed.). (1968-1975). Reading tests and reviews (2 vols.). Lincoln. NE: Buros Institute of Mental Measurements, University of Nebraska.

Bush, C.L., and Andrews, R.C. (1978). Dictionary of reading and learning disabilities. Los Angeles: Western Psychological Services.

Child Development Abstracts and Bibliography. (1927-). Lafayette, IN: Purdue University, Society for Research in Child Development.

Clark. B.R., and Neave. G.R. (Eds.). (1992). The encyclopedia of higher education (4 vols.). Tarrytown, NY: Pergamon Press.

Comprehensive Dissertation Index (1861-1972) (37 vols.). Ann Arbor, MI: University Microfilms.

Compton, C. (1984). .l guide to 75 diagnostic tests for special education. Belmont, CA: Fearon Education.

CompuServe Information Service Inc., 5000 Arlington Centre Boulevard, Columbus, OH 43220.

CTMG Publishers, 1501 Wilson Boulevard, Suite 1110, Arlington, VA 22209.

Current Index to Journals in Education. (1969-). Phoenix: Oryx Press.

Davis, W.E. (1980). Educator's resource guide to special education: Terms-laws-tests-organizations. Boston: Allyn & Bacon.

Dejnozka, E.L., and Kapel, D.E. (199 1). American educators' encyclopedia. Westport, CT: Greenwood Press.

Dialog Database Catalog. (1994). Palo Alto, CA: Dialog Information Services Inc.

Dialog Information Services Inc., 3460 Hillview Avenue, Palo Alto, CA 94304.

Digest of Educational Statistics. (1962-).Washington, DC: National Center for Educational Statistics.

Dissertation Abstracts International. (1938-). Ann Arbor, MI: University Microfilms.

Education Index. (1929-). New York: H.W. Wilson.

Educational Administration Abstracts. (1966-). Beverly Hills, CA: Sage Publications.

The Educational Testing Service Test Collection Catalog. (1986-). Phoenix: Oryx Press.

Eliot, J., and Macfarlane, I. (1983). An international directory of spatial tests. Atlantic Highlands, NJ: Humanities Press.

Exceptional Child Education Resources. (1969-). Reston, VA: Council for Exceptional Children.

Fustukjian, S.Y., and Taheri, B.J. (Eds.). (1992). Directory of ERIC information service providers. Washington, DC: Educational Resource Information Center (ERIC).

Gage, N.L. (Ed.). (1963). Handbook of research on teaching. Chicago: Rand McNally.

Good, C.V. (Ed.). (1973). Dictionary of education (3d ed.). New York: McGraw-Hill.

Hahn, H., and Stout, R. (1994). The Internet complete reference. Berkeley, CA: Osbourne McGraw-Hill.

Hardie, E.T., and Neou, V. (1994). Internet: Mailing lists. Englewood Cliffs, NJ: PTR Prentice Hall.

Higher Education Abstracts. (1984-). Claremont, CA: Claremont Institute for Administrative Studies.

Hill, PJ. (Ed.). (1982). A dictionary of education. Boston: Kegan Paul.

Houston, J.E. (Ed.). (1995). Thesaurus of ERIC descriptors (13th ed.). Phoenix: Oryx Press.

Husen, T., and Postlethwaite, TN. (Eds.). (1994). The international encyclopedia of education: Research and studies. Oxford: Pergamon Press.

Husen, T., and Postlethwaite, TN. (Eds.). (1985). The international encyclopedia of higher education (10 vols. and suppls.). Tarrytown, NY: Pergamon Press.

Index to International Statistics. (1982-). Bethesda, MD: Congressional Information Service.

Information Please Almanac. (1947-).Boston: Houghton Mifflin.

International Archives Institute, Inc. (InterArc), 105-B Douglas Court, Sterling, VA 22170.

Kehoe, B.P. (1994). Zen and the art of the Internet (3rd ed.). Englewood Cliffs, NJ: PTR Prentice Hall.

Keogh, B.K. (1980-). Advances in special education: A research annual. Greenwich, CT: JAI Press.

Keyser, D.J., and Sweetland, R.C. (Eds.). (1984-). Test critiques (Vol. 1-). Kansas City, M0: Test Corporation of America.

Knowles, A.S. (Ed.). (1978). The international encyclopedia of higher education (10 vols.). San Francisco: Jossey-Bass.

Krol, E. (1994). The whole Internet: User's guide & catalog (2nd ed.). Sebastopol, CA: O'Reilly & Associates Inc.

Masters Abstracts: Abstracts of Selected Masters Theses on Microfilm (1962-1985). Ann Arbor, MI: University Microfilms.

Masters Abstracts International (1986-). Ann Arbor, MI: University Microfilms.

Marcaccio, K. (Ed.). Gale directory of databases (2 vols.). Detroit: Gale Research.

Mauser, A.J. (1981). Assessing the learning disabled: Selected instruments (3d ed.). Novato, CA: Academic Therapy Publications.

Moore, B.C., Abraham, W., and Laing, C.R. (1980). A dictionary of special education terms. Springfield, IL: Thomas.

Motley, L. (1994). Modem GSA: Low cost and free on-line sources for information, the Internet, databases, and electronic bulletin boards via personal computer and modem in 50 states and Washington, D.C. (2d ed.). Takoma Park, MD: Allium Press.

Murphy, L.L., Conoley, J.C., and Impara, J.C. (Eds.). (1994). Tests in print IV. Lincoln, NE: Buros Institute of Mental Measurements, University of Nebraska-Lincoln.

O'Brien, JH., and Wasserman, S.R. (Eds.). (1994). Statistics sources (16th ed., 2 vols.) Detroit: Gale Research.

Page, G.T, Thomas, J.B., and Marshall, A.R. (1980). International dictionary of education. Cambridge, MA: MIT Press.

PsychologicalAbstracts (1927-). Lancaster, PA: American Psychological Association.

Resources in Education (1968-). Washington, DC: U.S. Government Printing Office.

Review of Educational Research (1931-). Washington, DC: American Educational Research Association.

Review of Research in Education (1973-). Washington, DC: American Educational Research Association.

Reynolds. C.R., and Mann, L. (Eds.). (1987). Encyclopedia ofspecial education: A reference for the education of the handicapped and other exceptional children and adults (3 vols.). New York: John Wiley.

Robinson, J.P., and Shaver, P.R. (1973). Measures ofsocial psychological attitudes. Ann Arbor, MI: Survey Research Center, Institute for Social Research.

Sheehy, E.P. (1986). Guide to reference books (10th ed.). Chicago: American Library Association.

SilverPlatter Information Inc., 100 River Ridge Drive, Norwood, MA 02062.

Silvey. H.M. (1951-). Master's theses in education. Cedar Falls, IA: Research Publications.

Social Sciences Citation Index. (1973-). Philadelphia: Institute for Scientific Information.

Social Sciences Index. (1974-). New York: H.W. Wilson.

Sociological Abstracts (1953-). San Diego, CA: Sociological Abstracts, Inc.

Sociology of Education Abstracts. (1965-). Abbington, England: Carfax Publishing.

Statistical Abstract of the United States. (1878-). Washington, DC: U.S. Bureau of the Census.

Statistical Reference Index. (1980-). Bethesda, MD: Congressional Information Service.

Sweetland, R.C., and Keyser, D.I. (Eds.). (1986). Tests: A comprehensive reference for assessments in psychology, education and business. Kansas City, M0: Test Corporation of America.

Tennant, R., Ober, J., and Lipow, A.G. (1993). Crossing the Internet threshold: An instructional handbook. Berkeley, CA: Library Solution Press.

Tests in Microfiche: Annotated Index. (1975-). Princeton: Educational Testing Service.

Travers, R. (Ed.). (1973). Second handbook of research on teaching (2d ed.). Chicago: Rand McNally.

UNESCO. (1984). Current bibliographical sources in education (2d ed.). Paris: UNESCO, International Bureau of Education.

U.S. Library of Congress, Exchange and Gift Division, (1910-). Monthly checklist of state publications. Washington, DC: U.S. Government Printing Office.

U.S. Superintendent of Documents. (1895-). Monthly catalog of US. government publications. Washington, DC: U.S. Government Printing Office.

Walberg, H.1., and Haertel, G.D. (Eds.). (1990). The international encyclopedia of educational evaluation. Elmsford, NY: Pergamon Press.

Waynar, B.S. (Ed.). (1970-).American reference books annual. Littleton, C0: Libraries Unlimited.

Wittrock, M.C. (Ed.). (1985). Handbook of research on teaching (3d ed.). New York: Macmillan,

Wolman, B.B. (1973). Handbook ofgeneral psychology. Englewood Cliffs. NJ: Prentice Hall.

Woodbury, M. (1982). A guide to sources of educational information (2d ed.). Arlington, VA: Information Resources Press.

WorldAlmanac and Book of Facts. (1868-). New York: Newspaper Enterprise Association.

الفصل الرابع

الفرضية

The Hypothesis

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون بوسع الطالب أن:

1. يعرف الفرضية.
2. يذكر أهداف الفرضية (أو الفرضيات) في دراسة بحثية ما.
3. يميز بين فرضية استقرائية وفرضية استنتاجية.
4. يذكر المعايير المستخدمة لتقييم فرضيات من أجل البحث.
5. يعرف (التعريف الإجرائي) ويعطي مثالاً عليه.
6. يحدد فرضية قابلة للاختبار من أمثلة معطاة.
7. يعرف (الفرضية الصفرية) ويفسر غاياتها في دراسة بحثية.
8. يكتب فرضية بحثية وفرضية صفرية من أجل دراسة بحثية.
9. يميز بين فرضية موجهة وأخرى غير موجهة.
10. يصف خطوات اختبار فرضية ما.
11. يذكر غاية (أو غايات) متوخاة من دراسة استطلاعية / تمهيدية.
12. يدون قائمة العناصر التي ستحويها خطة بحث ما.

بعد بيان مشكلة البحث وتفحص الأدبيات، يكون الباحث مهيئاً لصياغة الفرضية. والفرضية عبارة يتبناها الباحث لتفسير محتمل لظاهرة ما، ثم يبني عليها محاكمته المنطقية كما لو كانت الفرضية صحيحة وذلك في محاولة للوصول إلى حقيقة لم تكن معروفة مسبقاً. وهي تعرض توقعات الباحث حول العلاقة بين المتغيرات داخل مشكلة البحث.

ويجري بعد ذلك اختبار الفرضية في دراسة بحثية. ومن ثم يجري تقديمها على أساس كونها مجرد حل مقترح لمشكلة البحث، مع إدراك أن التمحيص اللاحق لها قد يؤدي إلى الاحتفاظ بها أو رفضها.

والفرضية أداة قوية في الاستقصاء العلمي. وهي تمكننا من ربط النظرية بالملاحظة، والملاحظة بالنظرية. وفي الوقت الحاضر، فإن الفرضيات تمكننا أثناء بحثنا عن المعرفة باستخدام آراء الفلاسفة الاستقرائيين مع تأكيدهم على المراقبة أو الملاحظة وكذلك باستخدام منطق الفلاسفة الاستنتاجين في تأكيدهم على التعليل المنطقي. وقد قامت الفرضيات بتوحيد الخبرة والتعليل المنطقي لإنتاج أدلة قوية تبحث عن الحقيقة.

فمثلاً، قد يبتدئ شخص ما بالسؤال: ما هو دور إدراك الأطفال لذواقم أثناء عملية تعلم القراءة؟ فقد يفترض المرء عندئذ، بأن هناك علاقة إيجابية بين إدراك الأطفال لذواقم وتحصيلهم في القراءة في السنة الأولى. أو قد يبتدئ شخص آخر بسؤال مثل: ما تأثير التدريب قبل دخول المدرسة على تحصيل أطفال محرومين ثقافياً في السنة الأولى؟ ومنطوق الفرضية يمكن أن يكون: تحصيل الأطفال المحرومين ثقافياً الذين توفر لهم تدريب قبل دخول المدرسة يكون أعلى خلال السنة الأولى من أطفال محرومين ثقافياً ولم يتوفر لهم تدريب قبل دخول المدرسة. ويمكن في كلا المثالين ملاحظة أن الفرضية هي مقترح يربط بين أثنين من المتغيرات. ففي الفرضية الأولى كان المتغيران هما ادراك الذات والتحصيل في القراءة، وفي الثانية هما وجود أو عدم وجود تدريب ما قبل المدرسة والتحصيل في السنة الأولى.

وبالرغم من الإشادة بالفرضيات وكونها تخدم عدة غايات مهمة، إلا أنها ليست ضرورية (بصورة مطلقة) في جميع الدراسات البحثية. فالفرضيات أدوات تستخدم أثناء عملية البحث، وليست أهدافاً في حد ذاتها. وغالباً ما تجري الدراسات في مجالات تكون فيها المعلومات الجوهرية المتجمعة ضئيلة. فإذا ما أفتقر باحث إلى إدراك عميق في مجال مشكلة بحثية ما، أو في متغيرات رئيسية تؤثر في إحدى الظواهر، أو في الطريقة التي تعمل فيها المتغيرات، فيكون عندئذ من الصعب جداً صياغة فرضية ذات معنى. فمثلاً، المسوحات التي تسعى لوصف ميزات ظاهرة معينة ما، أو للتأكد من مواقف وآراء مجموعات إنسانية، غالباً ما تمضي في سبيلها دون فرضيات. أما في البحث النوعي، فمن النادر صياغة فرضيات في بداية البحث. فالفرضيات عادة ما تتولد عندما تتجمع البيانات وعندما يحصل الباحث على عمق بصيرة أكبر وفهم للظاهرة تحت التمحيص.

وهناك سببان لصياغة فرضية ما قبل مرحلة تجميع بيانات أية دراسة وهما: (1) الفرضية القائمة على مرتكز قوي تشير إلى إن للباحث معرفة كافية للشروع في البحث، (2) الفرضية ترشد عملية تجميع وتفسير البيانات، كما أن الفرضية توضع للباحث السبيل الذي يتوجب عليه سلوكه، ونمط البيانات التي يتوجب جمعها، وبذلك يتجنب الباحث هدر الكثير من الوقت والجهد.

الأهداف الرئيسية التي تقدمها الفرضية تتضمن الآتي:

1- تقدم الفرضية تفسيراً تجريبياً (مؤقتاً) للظاهرة، كما تقوم بتيسير توسيع المعرفة في مجال ما. فالوصول إلى معرفة يمكن الاعتماد عليها حول المشكلات التربوية، يحتم علينا الذهاب أبعد من مجرد تجميع الحقائق المعزولة، وذلك بالبحث عن التعميمات والعلاقات القائمة بين تلك الحقائق. فتلك العلاقات والتعميمات تُهيئ التكيف المهم لفهم المشكلة. وليس محتملا أن يتضح مثل هذا التكيف طالما كان تجميع البيانات دون توجيه. ومن خلال توضيحات مقترحة، تُهيئ الفرضيات توجيهاً لمحاولة بحثية، كما يمنحها التركيز. وبسبب إمكانية اختبار الفرضيات وتأييد صحتها من خلال الاستقصاء العلمي، فأن الفرضيات تسمح لنا بتوسيع المعرفة.

2- تتيح الفرضية للباحث عبارة علاقية يمكن اختبارها مباشرة في دراسة بحثية. فلا يمكن اختبار الأسئلة مباشرة. ويبدأ البحث بسؤال، لكنه يمكن اختبار العلاقة المقترحة بين المتغيرات فقط . وعلى سبيل المثال، لا يمكن للمرء أن يختبر السؤال: (هل ينتج عن تعليقات المدرسين على أوراق الطلبة تقدم متميز في تحصيل الطلبة؟) فعوضاً عن ذلك، يختبر المرء الفرضية التي يتضمنها السؤال: (تعليقات المدرسين على أوراق الطلبة ينتج عنها تقدم متميز في التحصيل)، أو بطريقة محددة أكثر: درجات الأداء لطلبة تمتعوا بتعليقات مدرسيهم على أوراقهم السابقة سوف تزيد عن درجات أولئك الطلبة الذين لم تتوفر تعليقات مدرسيهم على أوراقهم السابقة. عندئذ، ينتقل المرء إلى بحث عن العلاقة بين المتغيرين (تعليقات المدرسين، وأداء الطلبة).

3- تهيئ الفرضية توجيها لنحو البحث. فالفرضية تقترح علاقة محددة بين المتغيرات، وبذلك يتم تقرير طبيعة البيانات اللازمة لاختبار القضية المطروحة. وببساطة شديدة، تقول الفرضية للباحث ماذا يعمل. ولابد من اختيار الحقائق وإجراء الملاحظات ذات الصلة بسؤال محدد، أما الفرضية فإنها تقرر صلة هذه الحقائق. وتُهيئ الفرضية أساساً لاختيار القياس المتعلق بانتقاء العينة وأساليب البحث التي ستستخدم وكذلك التحليل الإحصائي المناسب. وبالإضافة إلى ذلك، فإن الفرضية تساعد في إبقاء الدراسة محددة المجال، وتحول دون أن تغدو واسعة جداً أو غير عملية.

فمثلاً، دعونا نتأمل مرة أخرى الفرضية المتعلقة بالتدريب لفترة ما قبل المدرسة الابتدائية للأطفال المحرومين ثقافياً، وأداءهم في السنة الابتدائية الأولى. فهذه الفرضية تشير إلى الطريقة

اللازمة للبحث، والعينة التي ستستخدم، بل أنها تقوم بتوجيه الباحث إلى الاختبار الإحصائي الضروري لتحليل البيانات. ويتضح من بيان الفرضية أن الباحث سوف يقوم بدراسة عليّة- مقارنة، تقارن تحصيل عينة من السنة الأولى لأطفال محرومين ثقافياً ممن تدربوا قبل دخولهم المدرسة، ومجموعة مماثلة لأطفال لم يتوفر لهم تدريب قبل المدرسة. وأي اختلاف في متوسط التحصيل لكلا المجموعتين يمكن تحليله من أجل الدلالة الإحصائية باستخدام اختبار "ت" (t-test)، أو بأسلوب تحليل التباين (وستجرى مناقشة هذه الأساليب في الفصل السادس).

4. تقدم الفرضية إطاراً لنشر نتائج الدراسة. وسيجد الباحث إن من الملائم جداً أخذ كل فرضية على حدة، وصياغة النتائج الخاصة بها، وهذا معناه، أن الباحث يتمكن من ترتيب هذا الجزء من التقرير المكتوب حول ما وفرته الإجابات للفرضيات الأصلية، وبهذا يكون للعرض معنى أكبر وقابلية أكبر للقراءة.

اقتراحات لاشتقاق الفرضيات
SUGGESTIONS FOR DERIVING HYPOTHESES

كيف يشرع الباحث في اشتقاق فرضية ما؟ وكما هو موضح في الفصل الثاني، فإن أصل دراسة باحث ما يمكن أن تكون مشكلة عملية، أو في أحد المواقف السلوكية التي تحتاج إلى إيضاح، أو في بحث سابق، أو حتى بما هو أجدى في إحدى النظريات التربوية، أو النفسية أو الاجتماعية. وهكذا، فإن الفرضيات يجري اشتقاقها استقرائياً من ملاحظات للسلوك الإنساني أو استنتاجياً من نظرية ما، أو من نتائج بحوث سابقة.

الفرضيات الاستقرائية Inductive Hypotheses

يقوم الباحث في النهج الاستقرائي، بصياغة إحدى الفرضيات كتعميم ناتج عن ملاحظة العلاقات، وهذا معناه، قيام الباحث بملاحظات السلوك الإنساني، ومراقبة الاتجاهات أو العلاقات المحتملة، ويفترض، بعدئذ، أحد التفسيرات لذلك السلوك الملاحظ. وبالطبع، لابد أن يصاحب عملية الاستدلال هذه، فحص للبحوث السابقة لتحديد النتائج التي توصل إليها باحثون آخرون حول المسألة. والنهج الاستقرائي مصدر مثمر على وجه الخصوص للفرضيات بالنسبة لمدرسي الصفوف. فالمدرسون يقومون بمراقبة التعلم والسلوك الآخر للطلبة في كل يوم، ثم يحاولون ربط ذلك السلوك بسلوكهم الذاتي وبسلوك تلاميذ آخرين، وبطرق التعليم المستخدمة، وبالتغيرات في البيئة المدرسية، وهكذا. واستناداً إلى خبراتهم ومعلوماتهم عن السلوك في الظروف المدرسية، قد يقوم المدرسون استقرائياً بصياغة أحد التعميمات التي تسعى لتفسير العلاقة الملاحظة. ومع ذلك، لابد من تقرير صحة هذا التفسير، وبهذا يمكن له أن يصبح فرضية لبحث علمي.

وربما قام أحد المدرسين بملاحظة درجة عالية من القلق الناشئ عن اختبارات صفية، فيعتقد بأن لهذا تأثير ضار على أداء التلاميذ. والأكثر من ذلك، فإن ذلك المدرس قد لاحظ بأنه إذا ما أعطي التلاميذ فرصة لكتابة تعليقات حول أسئلة موضوعية، فأن أداءهم في الاختبار يبدو كأنه تحسن. ويستدل المدرس منطقياً بأن تلك الحرية في كتابة التعليقات لابد أن تساعد بطريقة ما في تخفيض القلق- وكنتيجة لذلك- ينال التلاميذ درجات أفضل. وتقترح تلك الملاحظة فرضية، هي: الطلبة الذين قد جرى تشجيعهم على كتابة التعليقات حول فقرات اختبارهم على أوراق الإجابات سوف يحرزون درجات أعلى من طلبة لم تتسنى لهم فرصة لكتابة أية تعليقات. وبوسع مدرسنا هذا أن يقوم بتخطيط تجربة لاختبار مدى صحة هذه الفرضية. لاحظ أن الفرضية تعبر عن اعتقاد المدرس بخصوص العلاقة بين المتغيرين وهما: (كتابة التعليقات حول فقرات اختبار، والأداء في الاختبار). لاحظ أيضاً، أن المتغير (القلق) الذي كان جزءاً من السلسلة الاستنتاجية المؤدية إلى الفرضية ليس جزءاً من الفرضية النهائية. بناء على ذلك، فإن نتائج البحث ستقدم معلومات تخص، فحسب، العلاقة بين كتابة التعليقات والأداء في الاختبار. أما العلاقات بين القلق والتعليقات، والقلق والأداء في الاختبار، فيمكن أن تكون موضوعات لفرضيات لاحقة يمكن بحثها. وغالباً ما يجد المرء أن إحدى الأفكار الأصلية تشمل سلسلة من العلاقات التي لا يمكن ملاحظتها بصورة مباشرة. وعندئذ، يقوم المرء بصياغة المسألة من اجل التركيز على العلاقات التي يمكن لها أن تخضع للملاحظة المباشرة.

وهذه بعض الأمثلة الأخرى على فرضيات قد تنتج عن ملاحظات المدرس.

- توجد علاقة إيجابية بين قلق (الرياضيات) وقلق (الحاسبات الإليكترونية) لدى تلامذة المدارس الثانوية.
- يزيد تعلم التلاميذ لبرمجة الحاسوب في المرحلة المتوسطة في نمو مهاراتهم في التفكير المنطقي.
- استخدام المنظم المتقدم/ التمهيدي(*) يزيد من تعلم تلاميذ المدارس الثانوية في الكيمياء باستخدام التعليم المعان بالحاسوب (Computer-assisted instruction)
- توجد علاقة إيجابية بين إدراك الذات الأكاديمي، وتوقعات النجاح، والتحصيل الأكاديمي.
- إن الطلبة المتدربين على كتابة ملخصات إحدى المحاضرات سوف يكون أداؤهم أفضل في اختبار بعدي فجائي، حول فهم المحاضرة، من الطلبة الذين يقومون فقط بكتابة الملحوظات.
- الدرجات التي يحصل عليها الأطفال في القراءة في اختبار نهاية السنة الأولى تكون في حالة المجموعات الصغيرة أعلى من حالة المجموعات الكبيرة.

(*) المنظم المتقدم/ التمهيدي (Advance Organizer): مفهوم عرضه أوزبل (Ausubel) في الستينات من القرن الماضي في نظريته "التعلم ذو المعنى" ويشير به إلى عبارة تمهيدية أو مناقشة أو أي نشاط آخر، تعرض فيه مادة جديدة على مستوى أعلى، من حيث العمومية والتجريد والشمول، من مهمة التعلم الفعلية الجديدة - (المراجع)

- يتأثر التطور الإدراكي والوجداني لأطفال السنة الأولى بكمية الخبرة التي يحصل عليها الأطفال قبل دخولهم المدرسة.

ففي العملية الاستقرائية، يقوم الباحث بالملاحظات، و يفكر في المشكلة ويعود إلى الأدبيات ليتزود بالقرائن، ثم يقوم بملاحظات إضافية، وبعد ذلك يقوم بصياغة إحدى الفرضيات التي تسعى لتفسير السلوك الملاحظ. ويجري بعدئذ اختبار الفرضية تحت ظروف منضبطة من أجل اختبار افتراض المدرس علمياً وهو المختص بالعلاقة بين المتغيرات المعنية. إن استقصاءات كهذه لفرضيات استقرائية مشتقة من الخبرات اليومية يمكن أن ينتج عنها فهم أكبر للعلاقات بين ممارسات المدرس داخل الصف وتعلم التلاميذ.

الفرضيات الاستنتاجية Deductive Hypotheses

خلافاً لفرضيات صيغت كتعميمات من علاقات تمت ملاحظتها، هناك فرضيات أخرى مشتقة بطريقة الاستنتاج من النظريات. وتقود هذه الفرضيات إلى نظام معرفي أكثر عمومية، كإطار يدمجها جميعاً بطريقة ذات معنى في مجموعة المعارف الموجود فعلاً ضمن النظرية ذاتها. ولا يمكن لأي علم أن يتطور بصورة كفوءة إذا ما بقيت كل دراسة لوحدها كجهد منعزل. ويصبح علماً تراكمياً بالبناء على مجموعة الحقائق والنظريات القائمة. وتعرف الفرضية المشتقة من إحدى النظريات بكونها فرضية استنتاجية.

وربما يكون، من المفيد، التمييز بين النظرية والفرضية. فالنظرية تتضمن مجموعة من المفاهيم، إضافة إلى عبارات حول الكيفية والأسباب المتعلقة بالعلاقات المتبادلة القائمة بين تلك المفاهيم. ويفترض بالنظريات أنها تقدم تفسيرات عامة يمكن تطبيقها على مجال واسع من الظواهر. وداخل إطار النظرية فإن العبارات ليست سوى فرضيات نعتبرها صحيحة - أي فرضيات يمكن اختبارها في دراسات لاحقة. ومن العلاقات المتبادلة المقترحة داخل النظرية، يمكن للمرء أن يتوصل إلى نتائج نفترض منطقيا أنها لازمة عنها. وهذه النتائج المفترضة تشكل أساساً للفرضيات. ويتحتم على النظرية العلمية أن تنطوي بالضروري على نتائج يمكن التحقق منها بالاستقصاء التجريبي، وهذا يعني: أنه يجب أن يكون المرء قادراً من النظرية على التنبؤ بملاحظة أو عدم ملاحظة أحداث معينة. وتصبح هذه النتائج المستنتجة فرضيات خاضعة إلى استقصاء تجريبي.

وعندما تتلقى الفرضيات المشتقة من إحدى النظريات دعماً في البحث العلمي، فإن النظرية بدورها أيضاً تتلقى دعماً. وهكذا، تقدم الفرضيات الدليل الذي يدعم، أو يوسع أو يناقض، أو يؤدي إلى تنقيح النظريات التي تم اشتقاق الفرضيات منها. وهنا يجب التأكيد على أنه في الوقت الذي يتمكن البحث العلمي من دحض إحدى النظريات، إلا أنه لا يمكنه البرهنة مطلقاً على صحة النظرية. ويعود ذلك إلى أن النظريات هي تعميمات تنطبق عملياً على جميع

الأمثلة/ الحالات الممكنة للظواهر التي تسعى النظريات لتفسيرها، وليس في الإمكان اختبار نظرية إزاء جميع الإمكانات. غير أنه، كلما زاد الدعم الذي تتلقاه النظرية في مجموعة متنوعة من الدراسات البحثية، زادت الثقة لدينا بأن النظرية صحيحة ومفيدة.

النظريات أكثر عمومية في المحتوى من الفرضيات، ويمكن للنظرية الواحدة يمكن أن تنتج عدة فرضيات لاستقصائها في عدة دراسات منفصلة. وربما يقوم شخص ما باستهلال دراسة ما وذلك باختيار إحدى النظريات في مجال اهتمامه. فالنظرية الخاصة التي يختارها المرء في بحثه، يحددها، بالطبع، هدف البحث وكذلك الإسهام الذي بوسع النظرية أن تقدمه لفهم المشكلة. وحالما يتم اختيار النظرية، يشرع المرء باشتقاق إحدى الفرضيات من النظرية باستخدام التفكير الاستنتاجي وذلك للوصول إلى النتائج المنطقية للنظرية. وتغزو هذه الاستنتاجات،بعدئذ، الفرضيات لتلك الدراسة البحثية.

فالباحث المهتم بالبحث في الاستيعاب القرائي لدى الأطفال قد يختار نظرية إدركية كنقطة البداية. وخلال السنوات القليلة الماضية، فإن البحث المرتكز على هذه النظرية قد قاد المربين إلى إعادة التفكير في المفاهيم السائدة، وإلى النظر إلى الاستيعاب القرائي بكونه عملية أكثر تعقيداً من مجرد اكتساب مجموعة من المهارات. وتؤكد النظرية الإدراكية/ المعرفية الطبيعة المتفاعلة للقراءة، والطبيعة البنائية للاستيعاب- وهذا معناه، أن القارئ يستخدم معرفته القائمة ويدمجها بمعرفة جديدة، ليبني بحيوية معنى من النص. وتنص النظرية على أنه بالإضافة إلى المعرفة، فإن لدى القارئ الجيد مجموعة من الاستراتيجيات المهمة المستخدمة في القراءة تقوم على استخراج الاستنتاجات لملء التفصيلات المحذوفة في النص.

وإذا ما أراد شخص ما اختبار مضامين هذه النظرية لتدريس الاستيعاب، فبوسعه استخدام التفكير الاستنتاجي ليصل إلى نتيجة منطقية للنظرية التي يمكن التحقق منها تجريبياً. والاستنتاج هو الفرضية في الدراسة البحثية. فمثلاً، يتمكن المرء أن يستنتج أن تعليم الأطفال استراتيجيات الاستدلال سوف ينتج عنه تحسن في الاستيعاب القرائي . ويمكن صياغة المشكلة البحثية والفرضية كالآتي:

المشكلة: ما تأثير تدريب وممارسة الاستدلال على الاستيعاب القرائي لدى الأطفال؟

الفرضية: إن الأطفال الذين تعلموا مهارات الاستدلال سيكون استيعابهم أكبر للنصوص من أطفال لا يتضمن تعليمهم مهارات الاستدلال.

سيشرع الباحث، بعدئذ، في تخطيط دراسة لاختبار الفرضية. وإذا ما تمت ملاحظة النتائج المتوقعة، عندئذ تتلقى النظرية الدعم. وفي الحقيقة، فإن هذه الفرضية قد جرى بحثها ودعمها في عدد من الدراسات (Hansen & Pearson, 1983). ومع ذلك، هناك حاجة إلى بحوث إضافية كثيرة، حول مضامين النظرية الإدراكية لتدريس القراءة.

وثمة نظرية أخرى مفيدة يمكن القيام باستنتاجات منها للبحث، وهي نظرية (بياجيه) حول تطور التفكير المنطقي لدى الأطفال. و قد اقترح (Piaget, 1968) بأن الأطفال يمرون عبر مراحل متنوعة أثناء تطورهم العقلي، وإحدى هذه المراحل هي مرحلة العمليات المادية/ المحسوسة، التي تبدأ في السنة السابعة أو الثامنة، وهي تتميز بالانتقال من الاعتماد على الإدراك الحسي إلى القدرة على استخدام بعض العمليات المنطقية. وهذه العمليات تقع على مستوى محسوس ، لكنها تشمل حقيقة بعض التفكير الرمزي. واستخدام هذه النظرية كنقطة انطلاق يجعل الباحث قادراً على افتراض أن نسبة الأطفال ذوي التاسعة من عمرهم والذين سوف يكونون قادرين على الإجابة الصحيحة على مشكلة الاستنتاج الانتقالي وهي: (فرانك أطول من جورج، وجورج أطول من روبرت. فمن هو الأطول؟)، أعلى من نسبة الأطفال في سن السادسة الذين يتمكنون من الإجابة الصحيحة على تلك المشكلة.

في الدراسة المخصصة لاختيار أحد الاستنتاجات من إحدى النظريات، من المهم جدا، التدقيق في إمكان وجود أية فجوات منطقية بين النظرية والفرضية. ويتحتم على الباحث السؤال: (هل تلزم الفرضية منطقيا عن النظرية؟) وان لم تلزم الفرضية عن النظرية، عندئذ لا يتمكن الباحث من الوصول إلى نتائج صحيحة حول كفاية النظرية. وإذ ما لاقت الفرضية دعماً، دون أن تكون مستنتجة بدقة من النظرية، فيتعذر على الباحث القول بأن نتائج البحث تضفي مصداقية على النظرية، وبالعكس، إذا لم تدعم البيانات الفرضية، فإن النظرية التي انبثقت منها أصلا، لن تكون بالضرورة أقل مصداقية.

صحيح إن الكثير من الفرضيات التي يمكن استنتاجها من نظريات معروفة قد جرى اختبارها فعلاً، لكن الكثير من تلك الاستنتاجات لا تزال تنتظر التوصل إليها، ومن ثم اختبارها. ويمكن استخدام الاستنتاج الذي جرى بحثه سابقاً لتوليد فرضيات في ظروف متنوعة أكثر لتوسيع تطبيق النظرية.

خصائص الفرضية الصالحة للاستخدام

CHARACTERISTICS OF THE USABLE HYPOTHESIS

بعد صياغة الفرضية بصورة تجريبية، وقبل محاولة أية اختبارات تجريبية حقيقية، يجب تقييم إمكانية استخدام الفرضية كأداة بحث. ويجب أن تلبي الفرضية معايير معينة للمقبولية. ولا يمكن الحكم على القيمة النهائية لأية فرضية قبل اختبارها تجريبيا، لكن هناك معايير معينة تميز الفرضيات القيمة، ويتحتم على الباحث استخدامها للحكم على كفاية الفرضية المقترحة.

يجب أن تكون للفرضية قوة تفسيرية

A Hypothesis Must Have Explanatory Power

يجب على الفرضية أن تعرض تفسيراً ممكناً لما تحاول أن توضحه وهذا معيار واضح لكنه مهم. ولتوضيح ذلك، أفرض أنك تريد أن تشغل سيارتك، ولكن ذلك لا يتم. فالفرضية القائلة بأن السيارة سوف لن تبدأ بالتشغيل لأنك تركت الماء جارياً في مغسلة الحمام، ليست تفسيراً ممكناً. أما الفرضية التي تقول بأن البطارية فارغة فهذا تفسير ممكن ويستحق الاختبار.

يجب أن تبين الفرضية العلاقة المتوقعة بين المتغيرات

A Hypothesis Must State The Expected Relationship Between Variables

يتحتم على الفرضية أن تخمن العلاقة بين اثنين أو أكثر من المتغيرات. ففي المثال الذي قدمناه، سوف يكون من غير المفيد القول: (بأن السيارة سوف لن تعمل، ولديها جهاز كهربائي)، إذ لا توجد علاقة معينة بين المتغيرات، ونتيجة لذلك، لا توجد هناك علاقة مقترحة لاختبارها. فالفرضية المفيدة تكون كالآتي: لن تعمل السيارة بسبب وجود خلل في الجهاز الكهربائي. وهذا المعيار قد يبدو جلياً على نحو واضح، لكن تأمل العبارة التالية: (إذا ما اختلف الأطفال في إدراك الذات، فإنهم سوف يختلفون من واحد لآخر في تحصيلهم للدراسات الاجتماعية). وتبدو هذه الجملة كفرضية إلى أن تلاحظ أنه لا يوجد هناك بيان بعلاقة متوقعة. ويمكن وصف العلاقة المتوقعة كالآتي: إدراك الذات الأعلى محتمل له أن يكون سابقاً لتحصيل أعلى في الدراسات الاجتماعية. وهذه الفرضية يمكن أن تكون كالآتي: (هناك علاقة إيجابية بين إدراك الذات والتحصيل في الدراسات الاجتماعية) وإذا ما كان هناك توقع للعكس، فهذا يعني: (إدراك الذات الأعلى يؤدي إلى تحصيل أقل في الدراسات الاجتماعية). عندئذ ستكون الفرضية كالآتي: (هناك علاقة سلبية بين إدراك الذات والتحصيل في الدراسات الاجتماعية). وكل منطوق من هذين الاثنين سوف يفي بهدف المعيار الثاني.

يجب أن تكون الفرضية صالحة للاختبار

A Hypothesis Must Be Testable

إن أهم الصفات المميزة للفرضية (الجيدة) هي صلاحيتها للاختبار. والفرضية الصالحة للاختبار يمكن التحقق من صلاحيتها، بمعنى أنه يمكن استخلاص الاستنتاجات، أو النتائج أو الاستدلالات من الفرضية بطريقة تجعل الملاحظة التجريبية ممكنة بحيث تدعم أو لا تدعم الفرضية ذاتها. وإذا ما كانت الفرضية صحيحة، فلابد عندئذ، أن تكون بعض النتائج المعينة المتوقعة واضحة. والفرضية الصالحة للاختبار تجعل الباحث قادراً على تحديد فيما إذا كانت تلك النتائج اللازمة

عنها قابلة للحدوث فعلاً. وخلافاً لذلك، سوف يكون من المستحيل تأكيد أو عدم تأكيد الفرضية. ففي مثالنا، الفرضية القائلة: (إخفاق السيارة في التشغيل هو عقاب لي بسبب الذنوب التي اقترفتها) غير صالحة للاختبار، على ما يبدو، في هذه الدنيا.

وهناك فرضيات كثيرة، أو اقتراحات غير صالحة للاختبار أساساً. فمثلاً، الفرضية القائلة: (إن خبرة التلميذ قبل دخوله المدرسة تدعم التكيف العام لهذا التلميذ) سيكون اختبارها صعباً بسبب الصعوبة في تعريف وقياس التكيف العام. والمثال الآخر على فرضية غير صالحة للاختبار هي: (استخدام العمل المتكرر في مدرسة الفنون يعوق الإبداع الفني عند الطفل) وفي هذه الحالة، سوف تقوم مشكلات في تحديد وقياس الإبداع الفني، وكذلك في اختيار المعايير لتقرير إذا ما حدثت إعاقة للإبداع.

ولكي تكون الفرضية صالحة للاختبار، ينبغي أن تربط بين المتغيرات التي يمكن قياسها. وإذا لم تتوفر وسيلة لقياس المتغيرات سيكون عندئذ من المستحيل جمع البيانات الضرورية لاختبار صحة الفرضية. وهذا الأمر لا يمكن التأكيد عليه بصورة قوية جداً. فما لم يكن بالإمكان التحديد بدقة مؤشرات كل متغير، وقياس هذه المتغيرات لاحقاً، عندئذ تكون الفرضية غير صالحة للاختبار.

وتدعى مؤشرات المتغيرات، التعريفات الإجرائية. والتعريف الإجرائي، كما أوضحنا سابقاً، هو ذلك الذي يعرّف متغيراً ببيان (الإجراءات) أو الأساليب الضرورية لقياس ذلك المتغير. فمثلاً، لاحظ هذه الفرضية: (هنالك علاقة إيجابية بين احترام الذات لدى الطفل وتحصيله القرائي في السنة الابتدائية الأولى). ومن أجل أن تفي هذه الفرضية بمعيار القبول، من الضروري تعريف المتغيرات إجرائياً. فيمكن تعريف احترام الذات بالدرجات التي يتم الحصول عليها بواسطة (مقياس احترام الذات :Coopersmith, 1981)، وتعريف التحصيل القرائي بالدرجات التي يتم الحصول عليها بواسطة: (اختبار كاليفورنيا في القراءة) أو بواسطة تقديرات المعلمين في السنة الابتدائية الأولى للتحصيل القرائي.

وأول اعتبار يعطي عند صياغة فرضية ما، هو التأكيد على إمكانية إعطاء المتغيرات (تعريفات إجرائية). تجنب استخدام المفاهيم البنائية التي يستحيل إيجاد قياسات مناسبة لها. فالمفاهيم البنائية، مثل: الإبداع، التسلّط، الديمقراطية، وما شابهها، قد أضحى لها معاني متباينة بحيث أن الاتفاق على التعريفات الإجرائية لهذه المفاهيم سيكون صعباً، إن لم يكن مستحيلاً. وتذكر بأنه يجب تعريف المتغيرات بسلوك قابل للملاحظة والتعرف.

ومن المهم تجنب العبارات القيمية في أية فرضية. فعبارة مثل: (إن وجود برنامج إرشادي في المدارس الابتدائية أمر مرغوب فيه)، لا يمكن استقصاؤها في دراسة بحثية. وعلى الرغم من ذلك، فإن المرء يمكن أن يختبر الفرضية القائلة: (إن تلاميذ الابتدائية الذين تتوفر لهم برنامج

إرشادي سوف يعبرون لفظاً بالرضى عن مدرستهم بأكبر من أولئك التلاميذ الذين لم يتوفر لهم الإرشاد). فالمرء يتمكن من قياس التعابير اللفظية عن الرضا، لكن إن كانت مرغوبة أم لا فذلك يعتبر حكماً قيمياً.

يجب أن تكون الفرضية متوافقة مع مجموعة المعارف القائمة
A Hypothesis Should Be Consistent With The Existing Body of Knowledge

يجب أن لا تعارض الفرضيات، فرضيات سابقة وطيدة ولا النظريات والقوانين. فالفرضية (سيارتي سوف لن تعمل لأن سائل البطارية قد تحول إلى ذهب) تفي بالمعايير الثلاثة الأولى، لكنها متناقضة مع ما هو معروف حول طبيعة المادة إلى درجة لا تجعل المرء يتابعها. لكن الفرضية (السيارة سوف لن تبدأ بالعمل بسبب إن السائل الموجود في البطارية قد تبخر إلى مستوى منخفض) تتفق مع المعارف السابقة، وبناء على ذلك فهي تستحق المتابعة. ويحتمل أن يكون من غير المفيد افتراض عدم وجود علاقة بين مفهوم الذات للمراهقات وللمراهقين وبين نسبة نموهم البدني بسبب رجحان الدليل الذي يدعم علاقة كهذه.

وفي تاريخ العلم وجد بأن أناساً علماء مثل آينشتاين، نيوتن، داروين وكوبرينكوس قد طوروا، حقاً، فرضيات ثورية تناقضت مع ما كان مقبولاً من المعارف في زمانهم. لكن يجب أن نتذكر بأن العمل الذي قام به رواد كهؤلاء، لم يكن في واقعه نكراناً لمعارف سابقة، لكنه كان إعادة لترتيب المعارف في نظرية أكثر قبولاً. وفي أغلب الحالات، وخصوصاً بالنسبة للباحث المبتدئ، فمن الأسلم، الاقتراح بأنه ينبغي أن تكون الفرضية متفقة مع المعارف الموجودة في الساحة. ومرة أخرى، فإن هذا الأمر يشير إلى ضرورة القيام بمراجعة مقننة للأدبيات، كي تتم صياغة الفرضيات على أساس البحوث المنشورة، سابقاً، في المجال.

يجب أن يكون منطوق الفرضية بسيطاً وموجزاً بقدر الإمكان
A Hypothesis Should be stated as simply and concisely as possible

إن صياغة الفرضية بطريقة بسيطة لا يجعل اختبارها أسهل، فحسب، لكنه يهيئ أساساً لكتابة تقرير مفهوم وسهل في نهاية الدراسة. وغالباً ما يكون ضرورياً تقسيم فرضية عامة وواسعة المجال إلى عدة فرضيات محددة، لإتاحة إمكانية الاختبار والوضوح. فمثلاً، قام (Hines, Cruickshank, & Kennedy, 1985) باستقصاء العلاقة بين وضوح المدرس وتحصيل التلميذ ورضاه. وللكشف عن جميع جوانب هذا السؤال، قام الباحثون بصياغة أربع فرضيات: (1) هناك علاقة إيجابية ذات ارتباط عال بين مقاييس الوضوح المستند إلى التجربة وتحصيل المتعلم. (2) هناك علاقة إيجابية ذات ارتباط عال بين مقاييس الوضوح المستند إلى التجربة ورضا المتعلم.

(3) إدراك/ تصور التلاميذ لوضوح المعلم يتوسط/ يتدخل في تأثيرات الوضوح على التحصيل.
(4) إدراك/ تصور التلاميذ لوضوح المعلم يتوسط/ يتدخل في تأثيرات الوضوح على رضا المتعلم. وشرع الباحثون لتخطيط بحوث لاختبار كل من هذه الفرضيات. وقد وجدوا دعماً قوياً للفرضيات الأولى والثانية والرابعة، لكنه، كان هناك دعم أقل بكثير للعلاقة بين إدراك/ تصور وضوح المعلم والتحصيل.

وعموماً، فمن المقبول أن يقوم الباحث بصياغة إحدى الفرضيات لكل وجه ثانوي من أوجه مشكلة البحث، أو لكل وسيلة من وسائل جمع البيانات المستخدمة. فمثلاً، يمكن لباحث أن يبدأ بالفرضية التالية: (التلاميذ الذين تعلموا مادة الرياضيات باستخدام رزمة التعليم المعان بالحاسوب سيكون تعلمهم للمفاهيم الرياضية واحتفاظهم بها أكبر بكثير من أولئك التلاميذ الذين يستخدمون كتباً تقليدية مقررة).

ولما كان ضرورياً نشر نتائج كل من التعلم والاحتفاظ بالتعلم، فينبغي إعادة صياغة الفرضية الأصلية في فرضيتين منفصلتين. وستكون هاتان الفرضيتان كالآتي: 1) التلاميذ الذين يتعلمون الرياضيات باستخدام رزمة التعليم المعان بالحاسوب، سيكون تعلمهم للمفاهيم الرياضية أكبر بكثير من أولئك التلاميذ الذين يستخدمون كتباً مقرّرة تقليدية. 2) التلاميذ الذين يتعلمون الرياضيات باستخدام رزمة التعليم المعان بالحاسوب، سيكون احتفاظهم بالمفاهيم الرياضية أكبر بكثير من أولئك التلاميذ الذين يستخدمون كتباً مقررة تقليدية. وبهذه الطريقة، بالوسع إيضاح ما إذا كانت البيانات تدعم كل وجه خاص للمسألة العامة. وربما تقوم البيانات بالإشارة إلى فعالية التعلم باستخدام التعليم المعان بالحاسوب على التعليم الأصلي، وليس على الاحتفاظ به. ويتحتم أن لا يترعج المرء بسبب الإسهاب الكلامي الواضح في صياغة فرضيات مركبة، إذ يجب أن لا ننسى بأن زيادة الفرضيات المحددة أكثر نفعاً لأهداف قابلة للاختبار والوضوح.

ومن المفيد أيضاً أن تكون المصطلحات المستخدمة في الفرضية، أبسط ما يكون بحيث تكون مقبولة لنقل المعنى المقصود، كما يجب الابتعاد عن استخدام المفاهيم البنائية الغامضة. فالمطلوب استخدام مصطلحات بطريقة مقبولة عموماً لدى الإشارة إلى الظاهرة. وإذا كان هناك لاثنتين من الفرضيات قوة تفسيرية متماثلة، فالأبسط منهما هي المفضلة، لأنها ستقدم التفسير الضروري بافتراضات أقل ومتغيرات أقل لتعريفها. وتذكر أن مبدأ اقتصاد الجهد مهم في تقييم الفرضيات.

ويجري رفض الكثير من الفرضيات التي تمت صياغتها، وذلك بعد اختبارها تجريبيا. ونجد في تاريخ البحث العلمي بأن الفرضيات التي أخفقت في الحصول على الدعم فاقت بكثير في عددها تلك التي لاقت الدعم. ويدرك الباحثون من ذوي الخبرة بأن الفرضيات التي لم تجد التأييد تمثل جزءاً مفيداً و متوقعاً للخبرة العلمية. ويمكن لهذه الفرضيات أن تقود إلى إعادة النظر

في النظرية، وغالباً ما تقربنا من التفسير الصحيح للأمور. وقد كتب العالم (Darwin, 1989) قائلاً: "لقد حاولت دائماً أن أحافظ على تفكيري طليقاً من أجل التنازل عن أية فرضية مهما تكن أثيرة لديّ، (فأنا لا أتمكن من مقاومة صياغة فرضية ما في كل موضوع)، حالما تظهر الحقائق أنها تناقضها. ولم يكن لدي خيار سوى القيام بذلك، ، ولا أتذكر أية فرضية تمت صياغتها أول مرة ولم يجر التنازل عنها أو تحويرها كثيراً، عدا تلك المتعلقة بالجرف المرجاني". وحتى الفرضية غير المدعمة قد تكون مفيدة في كونها تشير إلى الحاجة إلى التأمل في الأوجه الأخرى للمشكلة، ويمكن بهذا أن تقرب الباحث خطوة أقرب لتفسير مقبول. ويتحتم على المرء لدى صياغة فرضية، أن ينصب جل اهتمامه على الابتعاد عن الغموض أو الالتباس.

وبالرغم من إيجاد الدعم لفرضية ما، فإن هذا لا يبرهن على صحة الفرضية، عدا حالات الاستقراء التام. ولا يمكن إطلاقاً البرهنة أو عدم البرهنة على صحة الفرضية. فالفرضية إما أن تكون مدعمة أو لا تكون. فالفرضيات بجورها ذات طبيعة احتمالية، فالدليل التجريبي يمكن أن يدفع المرء إلى الاستنتاج بأن من المحتمل أن يكون التفسير صحيحاً، أو أن من المعقول منطقياً قبول فرضية، لكن هذا لا يبرهن على صحة الفرضية مطلقاً. وفيما يأتي مثال على دراسة ذات فرضيات تم استنتاجها من إحدى النظريات (Davis, 1988).

تأثير وقت التدريس أثناء النهار على التحصيل
The Effect of Time of Day of Instruction on Achievement

مشكلة البحث Research Problem

ما تأثير وقت التدريس أثناء النهار، على تحصيل تلاميذ السنة الثامنة في مادتي الإنجليزية والرياضيات؟

النظرية Theory

هناك نظرية تتضمن الذاكرة قصيرة الأمد، والذاكرة طويلة الأمد(*) وتقول النظرية بأن الواجبات التي تشمل عموماً، استخدام الذاكرة قصيرة الأمل، يجري تعلمها وأداؤها بصورة أفضل في الصباح. والواجبات التي تشمل عموماً، استخدام الذاكرة طويلة الأمد يجري تعلمها وأداؤها بصورة أفضل بعد الظهر. وهنالك عاملان فسيولوجيان مسئولان عن الاختلافات في التعلم أثناء النهار: أ) إن التيقظ الجوهري يرتفع من مستوى منخفض صباحاً إلى الذروة مساءاً، واستخدام الذاكرة قصيرة المدى والذاكرة

(*) الذاكرة قصيرة الأمد (STM) وهي الذاكرة كما تقاس خلال خمس دقائق بعد استقبال المادة المراد تذكرها. أما الذاكرة طويلة الأمد (LTM)، فهي الذاكرة كما تقاس بعد مضي خمس دقائق على استقبال المثير كحد أدنى، وهي تقابل الذاكرة قصيرة الأمد (د. محمد علي الخولي: قاموس التربية) – المراجع.

طويلة المدى تترافقان معه بالترتيب. ب) أما الاتزان في مستويات البلازما والهرمونات فيؤثر في طريقة قيام الناس في تحويل المعلومات إلى لغة رمزية وخزنها واستعادتها، وهذه تترافق على نحو أفضل، مع معالجة الذاكرة طويلة المدى في وقت متأخر من النهار. ومع معالجة الذاكرة قصيرة المدى في وقت مبكر من النهار.

الفرضية Hypothesis

نظراً لأن القراءة تشتمل على ربط المعلومات المكتوبة بالمعارف السابقة والخبرات، فقد تم افتراض: (سيكون تحصيل الطلبة في القراءة أكبر عندما يكون التعليم بعد الظهر). ونظراً لأن تعليم المهارات الرياضية يشمل معالجات ذهنية للبيانات واستخدام الذاكرة قصيرة الأمد، فقد تم افتراض: (سيكون تحصيل الطلبة في الحساب عندما يكون التعليم صباحاً أكبر من حاله بعد الظهر).

الطرق Methods

لقد تم اختيار ثمانين طالباً بصورة عشوائية من السنة الثامنة في مجموعات المعالجة: 1) درس الحصة الأولى: الإنكليزية. 2) الحصة الأخيرة سوف تكون الإنجليزية. 3) الحصة الأولى: رياضيات. 4) الحصة الأخيرة: رياضيات. قام مدرس الإنجليزية ذاته ومدرس الرياضيات ذاته بتعليم الساعة الأولى والساعة الأخيرة للمجموعات، مستخدماً الأساليب والمواد ذاتها. تم تطبيق اختبار المهارات الأساسية الشامل (CTBS) في بداية السنة الدراسية (اختبار قبلي) وفي نهاية السنة الدراسية (اختبار بعدي) لقياس التحصيل. وأجريت الاختبارات في اليوم ذاته في كل جزء من المحتوي، وأثناء منتصف اليوم الدراسي لكل المجموعات من أجل التحكم في تأثير الوقت من النهار للمؤثرات الناتجة.

التحليل Analysis

لقد تم تحليل مجموعات درجات الإنجليزية والرياضيات ذات الصلة باستخدام التحليل المتغاير (ANCOVA) حيث درجات الاختبار القبلي توضح التغاير. وهذا التحليل يأخذ في حسابه التفاوت القائم بين درجات الطلبة في الاختيار القبلي، وينظم درجات الاختبار البعدي بالنسبة لهم، وذلك حتى تكون المقارنة صحيحة.

النتائج Results

حصلت مجموعة مادة الإنجليزية لفترة ما بعد الظهر على مستوى ذات دلالة أعلى من مجموعة مادة الإنجليزية لفترة الصباح. ولم يكن هناك فرق مهم ذو دلالة في تحصيل مادة الرياضيات بين مجموعتي الصباح وبعد الظهر.

الاستنتاجات Conclusions

سيستفيد تلاميذ السنة الثامنة من التعليم بعد الظهر أكثر من التعليم الصباحي في مادة الإنجليزي. لم يكن للوقت اليومي أية فروقات في تحصيل تلاميذ السنة الثامنة في الرياضيات. ولقد استنتج الباحث بأنه يجب للفرضية المتعلقة بالرياضيات أن تخضع إلى بحث إضافي. ولربما كانت وظائف الذاكرتين قصيرة وطويلة المدى التي وجدتها البحوث

السابقة لا تمثل كثيراً ما يتطلبه التفكير المنطقي الرياضي، أو بما يتضمن التفكير المنطقي الرياضي استخداماً متوازناً إلى حد لكلا الذاكرتين.

أنواع الفرضيات TYPES OF HYPOTHESIS

فرضية البحث Research Hypothesis

إن الفرضيات التي قمنا بمناقشتها حتى الآن تدعى فرضيات (البحث) أو الفرضيات الأساسية، وهي فرضيات تطورت عن الملاحظة، أو الأدب ذي الصلة، أو عن نظرية تم عرضها في الدراسة. وفرضية البحث هي بيان حول العلاقة التي يتوقع المرء إيجادها نتيجة للبحث. ويمكن أن تكون فرضية البحث عبارة عن (العلاقة) المتوقعة أو (الفرق) المتوقع بين متغيرات الدراسة. إن فرضية ما حول معامل ذكاء الأطفال (IQ) وقلقهم في غرفة الصف يمكن أن تكون كالآتي: (هناك علاقة إيجابية بين معامل ذكاء الأطفال وقلقهم في المدارس الابتدائية، أو: الأطفال الذين معامل ذكائهم عال سوف يكونون أكثر قلقاً في غرفة الصف من أطفال معامل ذكائهم متدن). ويمكن صياغة فرضيات البحث بشكل موجه أو غير موجه. فالفرضية الموجهة تحدد التوقع لطبيعة العلاقة أو الاختلاف. فالفرضيتان أعلاه حول معامل الذكاء والقلق موجهتان. والفرضية غير الموجهة، من جهة أخرى، تقول بأن العلاقة أو الفرق موجودان دون تحديد لطبيعة النتائج المتوقعة، مثل: هنالك علاقة بين معامل الذكاء والقلق عند الأطفال.

الفرضية الصفرية Null Hypothesis

ليس بوسعنا اختبار فرضيات البحث مباشرة بالوسائل الإحصائية المتاحة. ففي البحوث التجريبية حيث لابد من استخدام الاختبارات الإحصائية، يتحتم علينا ترجمة فرضية البحث إلى نوع آخر من الفرضيات يعرف (بالفرضية الصفرية H_0:). وتدعى صفرية إذ أنها تقول (ليس هناك فرق) أو (ليس هناك تأثير) أو (ليس هناك علاقة) فالفرضية الصفرية تنفي ما يتوقعه الباحث أو يتنبأ به. وربما يأمل أحد الباحثين بعد قيامه بإحدى المعالجات، بأن المتوسطين لمجتمعين إحصائيين مختلفان، لكن الفرضية الصفرية ستقول بأنه لا يوجد اختلاف بين متوسطي المجتمعين.

ويجري استخدام الفرضية الصفرية إذ أنها تمكن الباحثين من مقارنة نتائج بحوثهم مع توقعات الصدفة من خلال الاختبارات الإحصائية. فالفرضية الصفرية تفترض بأن الفروق الملحوظة حدثت بسبب الصدفة وحدها، ولهذا فهي لا تمثل فروقاً حقيقية أبداً. وتستخدم الاختبارات الإحصائية لتقرير احتمال صحة الفرضية الصفرية. وإذا ما دلت تلك الاختبارات بأن ثمة احتمالاً ضئيلاً لحدوث الفروق بالصدفة، عندئذ تصبح الفرضية الصفرية افتراضاً غير محتمل إلى حد بعيد. وعليه، يمكن رفضها لصالح فرضية بديلة. والدليل كاف للاستنتاج تجريبياً

بأن الاختلاف حقيقي. وإذا ما كانت الفروق الملحوظة تعود ببساطة إلى الصدفة، فالدليل غير كاف، و يجري الاحتفاظ بالفرضية الصفرية. أن اختبار فرضية صفرية يماثل عمل المدعي العام خلال محاكمة جنائية. فمن أجل إثبات الذنب، يتحتم على المدعي العام (في النظام الدستوري الأمريكي) تقديم الدلائل الكافية لتمكين المحلفين من رفض احتمال البراءة دون شكوك معقولة. وليس من الممكن لمدع عام البرهنة على وجود الذنب بطريقة حاسمة، وكذلك لا يتمكن باحث من الحصول على دعم غير مشكوك فيه لفرضية بحثية. فيفترض أن يكون المتهم بريئاً إلى أن توجد دلائل كافية تشير إلى غير ذلك، ويفترض أن تكون الفرضية الصفرية صحيحة إلى أن توجد دلائل كافية تشير إلى عكس ذلك.

فلنعد الآن إلى الدراسة الافتراضية لمستويات القلق لدى الأطفال ذوي معامل الذكاء المرتفع والمتدني. فقد نصوغ فرضية البحث كالآتي: (الأطفال ذوي معامل الذكاء المرتفع سيكونون أكثر قلقاً من الأطفال ذوي معامل الذكاء المتدني، بينما تقول الفرضية الصفرية: (إن مستوى القلق لا يختلف عن مستوى القلق لدى الأطفال ذوي الذكاء المتدني). وبصورة رمزية تصبح الفرضية الصفرية $H_0: \mu_1-\mu_2=0$. حيث μ_1: درجة متوسط القلق لمجتمع الأطفال ذوي معامل الذكاء المرتفع، μ_2 درجة متوسط القلق لمجتمع الأطفال ذوي معامل الذكاء المتدني.

ويجري اختبار الفرضية الصفرية باستخدام إحدى التقنيات الإحصائية المتنوعة التي سوف تجري مناقشتها في الفصل السادس. و إذا ما أشارت البيانات التجريبية بأن الاختلاف بين مجموعات العينة كان كبيراً إلى درجة لا يحتمل أن يكون بسبب الصدفة، عندئذ يمكن رفض الفرضية الصفرية. ورفض الفرضية الصفرية يشير إلى أن مستوى القلق لدى المجموعتين من الأطفال لا يحتمل أن يكون متساوياً. عندئذ يتبنى الباحث فرضية بديلة (نرمز لها H_1) تقول: (درجة متوسط القلق لدى المجموعتين غير متساو) "$H_1: \mu_1-\mu_2 \neq 0$"ويقال عندئذ بأن الفرضية البديلة غير موجهة، إذ أنها لا تبين المجموعة التي درجة متوسطها سيكون أكبر. وربما يبتدئ أغلب البحوث في مجال التربية بفرضيات غير موجهة. وفي بعض المواقف، ربما يكون لدى الباحث سبب معين لصياغة إحدى الفرضيات الموجهة، والتي كما يقول أسمها تحدد اتجاه نتائج البحث المتوقعة. والآتي هو مثال على فرضية موجهة: (متوسط درجات القلق لأطفال ذوي معامل ذكاء عال سيكون أكبر من متوسط درجات القلق لأطفال ذوي معامل ذكاء منخفض).

اختبار الفرضية TESTING THE HYPOTHESIS

إن أية دراسة تبتدئ بفرضية بحثية يتوجب فيها أن تكون عبارة بسيطة واضحة للعلاقة المتوقعة بين المتغيرات. وقد أوضحنا سابقاً بأنه يجب أن تكون الفرضيات صالحة للاختبار، وذلك يعني أنها طيّعة لتحقيق تجربي. وعندما يتكلم الباحثون عن اختبار فرضية ما، فأنهم يشيرون في ذلك

إلى الفرضية الصفرية. وفقط، الفرضية الصفرية هي التي يمكن اختبارها مباشرة بواسطة إجراءات إحصائية. واختبار أية فرضية يشمل الخطوات التالية.

1- بين، بمصطلحات إجرائية، العلاقات التي يجب ملاحظتها إذا ما كانت فرضية البحث صحيحة.
2- قم بصياغة الفرضية الصفرية.
3- اختر الطريقة التي ستتيح الملاحظة أو التجربة الضرورية لتوضح ما إذا كانت هذه العلاقات موجودة.
4- أجمع وحلل البيانات التجريبية.
5- حدد فيما إذا كان الدليل كافياً لرفض الفرضية الصفرية.

مثال على اختبار إحدى الفرضيات
An Example of Testing A Hypothesis

ربما يساعد أحد الأمثلة على توضيح أفضل لعملية اختبار فرضية تجريبياً. افرض بأن هناك أحد المعلمين المهتمين باستقصاء نظرية التعزيز داخل غرفة الصف. وقد يفترض المعلم، حسب فهمه لنظرية التعُزيز، أنه سينتج عن المديح أو التشجيع دافعية عالية لدى الطلبة، وسيقود بدوره إلى تحصيل أعظم. فإذا كانت هذه الفرضية صحيحة، فمن المنطقي افتراض بأن تعليقات المعلمين بالمديح على أوراق الاختبار سوف يتبعها تحسن في أدائهم.

(**الخطوة الأولى**): إن هذا المعنى الضمني يمكن صياغته كالآتي: تعليقات المدرسين على أوراق الطلبة ينتج عنها تحسن في أداء الطلبة في الاختبارات. أنها العلاقة بين المتغيرين: تعليقات المعلمين وأداء التلاميذ، وهو ما سيجري استقصاؤه.

(**الخطوة الثانية**): من أجل الاختبار الإحصائي، لابد من تحويل فرضية البحث أعلاه إلى فرضية صفرية. فالفرضية الصفرية تقول: تعليقات المدرسين على أوراق التلاميذ لا ينتج عنها تحسن في أداء التلاميذ لاختبارات.

(**الخطوة الثالثة**): يمكن بحث هذا النوع من الفرضيات بواسطة إجراء إحدى التجارب. وبوسع الباحث اختيار عدد من الصفوف الدراسية عشوائيا لاستخدامها في الدراسة. وداخل كل صف من الصفوف سيجري تقسيم التلاميذ عشوائياً إلى مجموعتين. فبالنسبة لطلبة المجموعة (أ) سيقوم مدرسوهم بكتابة تعليقات مشجعة تتعلق بأدائهم في الاختبار. [وهذه التعليقات سوف تكون ببساطة كلمات مشجعة للتلاميذ، مثل (ممتاز)، (تابع هذا العمل الجيد) ، (أنك أحسن الآن)]. (ويجب أن لا تكون لهذه التعليقات علاقة بمحتوى المادة أو بتصحيح الأخطاء لتلميذ معين، وإلا قد يُعزى التحسن إلى الفائدة التربوية لتعليقات كهذه وليس لزيادة الدافعية). أما طلبة المجموعة (ب) فلن يتلقوا أية تعليقات أبداً على أوراق اختبارهم.

(**الخطوة الرابعة**): سيقوم المدرسون بإجراء اختبار موضوعي يغطي وحدة معينة من المحتوى. وسوف يجري تصحيح الاختبارات، ويطبق على المعالجة التجريبية ما ذكر أعلاه. وبعدها، يجري المدرسون اختباراً ثان لوحدة تماثل في صعوبتها الوحدة الأولى حيث يكون قد تم تدريسها للتلاميذ بعد الاختبار الأول والمعالجة التجريبية. وسيتم التأكد من التغيير الذي طرأ من الاختبار الأول إلى الاختبار الثاني لدى كل تلميذ، وكذلك التحسن في معدل كل مجموعة. عندئذ، سوف يكون في الإمكان من خلال تحليل البيانات تحديد فيما إذا كان التحسن في المعدل في الاختبار الثاني عائداً للمعالجة التجريبية (وهي تعليقات المدرسين على الأوراق).

(**الخطوة الخامسة**): فإذا ما وجد بأن التلاميذ - كمجموعة واحدة - الذين تلقوا التعليقات (المجموعة أ) قد حصلوا على تحسن دال إحصائياً أعلى من المجموعة التي لم تتلق تعليقات (المجموعة ب)، عندئذ يمكن رفض الفرضية الصفرية. إن رفض الفرضية الصفرية سيعني بأن نتائج الطريقتين ليست متماثلة. وسوف يستنتج الباحث تجريبياً بأن تعليقات المدرسين على أوراق التلاميذ يؤدي إلى تحسن في أداء التلاميذ في الاختبارات.

دراسة تمهيدية / استطلاعية Pilot Study

قبل تهيئة وتحضير خطة البحث، ربما يكون من المفيد تجربة الإجراءات المقترحة على عدد قليل من الأفراد. وهذه المحاولة التجريبية أو الدراسة الاستطلاعية ستساعد، قبل كل شيء، الباحث على تقرير فيما إذا كانت الدراسة ممكنة التطبيق، وفيما إذا كانت تستحق الاستمرار فيها. وهي تهيئ فرصة لتقييم ملاءمتها وعمليتها لأدوات جميع البيانات. كما أنها تسمح بإجراء اختبار أولي للفرضية، بما قد يعطي بعض المؤشرات على مناعتها، كما تقترح فيما إذا كان الموقف بحاجة إلى تنقيح إضافي.

وستبين الدراسة الاستطلاعية كذلك كفاية إجراءات البحث وكفاية المقاييس التي تم اختيارها من أجل المتغيرات. وربما يمكن حل المشكلات التي لم يجر توقعها في هذه المرحلة، وبهذا يتوفر اقتصاد في الزمن والجهد فيما بعد. فالدراسة الاستطلاعية تستحق الوقت اللازم لها، كما تجري عادة التوصية بها للباحث المبتدئ.

THE RESEARCH PLAN خطة البحث

يكون المرء بعد صياغة المسألة والفرضية، جاهزاً لإكمال الخطة التجريبية للبحث. ويحتاج المرء إلى كتابة ما يقترح القيام به بالتفصيل، وما يخطط للقيام به فحسب.

إن تطوير خطة البحث أمر حيوي. ويضطر المرء إلى عرض أفكاره بشكل متماسك. فالكثير من الأفكار الأولية تبدو واعدة إلى أن يقوم المرء باكتشاف الصالح وغير الصالح منها،

وعندئذ تصبح الصعوبات أو القصور أموراً واضحة. ويمكن أيضاً إعطاء خطة البحث مكتوبة للآخرين من اجل تعليقهم ونقدهم. ومن السهل بمكان للآخرين، اكتشاف عيوب أو أخطاء مقترح ما، عندما يكون مكتوباً على الورق وليس عن طريق السماع الشفوي.

وبصورة نموذجية فأن خطة البحث عند هذه المرحلة هي عبارة عن مقترح أولي فقط، وربما تكون هناك حاجة إلى إجراء عدة تغييرات قبل كتابة المقترح الأخير بصورة رسمية. ومع ذلك، فإن من المفيد أن يتذكر المرء أنه كلما كان المقترح الأولي كاملاً ومفصلاً، كان أكثر فائدة للباحث، وللوقت الذي سيتوفر لاحقاً.

وتشمل عناصر خطة البحث عبارة المشكلة، والفرضية، وتصميم البحث، والعينة، والتحليل الإحصائي.

المشكلة The Problem

تبدأ خطة البحث بعبارة واضحة لمشكلة البحث. وتسأل العبارة عن العلاقة بين المتغيرات كما تحدد المجتمع ذا الأهمية. ويجب أن يكون هناك أيضاً وصف وجيز في هذه الفقرة، لخلفية مشكلة البحث في النظرية، وفي البحوث ذات الصلة.

الفرضية The Hypothesis

يتعين أن يعقب المسألة بيان وجيز للفرضية أو للفرضيات التي سيجري اختبارها. والفرضية تعطى اتجاهاً للبحث. وتعتمد جميع الخطط اللاحقة لمشروع البحث على منطوق الفرضية. ومن الواجب أن يقوم الباحث بكتابة الفرضية والأساس المنطقي لها بكل تحديد ووضوح. ويجب أن تشمل هذه الفقرة من خطة البحث التعاريف الإجرائية للمتغيرات المستخدمة.

تصميم البحث The Research Design

تقدم الفقرة التالية من الخطة وصفاً لتصميم البحث – وهذا يعني، وصفاً للإجراءات التي ستتبع في اختبار الفرضيات. ومن المهم جداً بأن يجري اختيار طريقة اختبار ملائمة. والسؤال التجريبي لا يمكن إجابته باستخدام وسائل مسح أو أية أساليب لا تحقق التقدم نحو النتائج المرجوة. فمثلا إذا قام باحث بتقييم مقارن لمزايا الطريقة الصوتية(*) والطريقة البصرية للقراءة(**) في مستهل تعليم القراءة، فقد يقترح الباحث القيام بمسح اتجاهات معلمي الابتدائية. إن مسحاً كهذا سوف لن يعطي جواباً ذا معنى يتعلق بالميزات النسبية للنظامين، إذ يقدم فقط آراء عينة من معلمين قد

(*) الطريقة الصوتية: طريقة في تعليم القراءة للمبتدئين تعتمد على أصوات الحروف لتنتقل من الحروف إلى الكلمة. (د. محمد على الخولي: قاموس التربية) – (المراجع)

(**) الطريقة البصرية: طريقة انظر وقل في تعليم القراءة. (د. محمد على الخولي: قاموس التربية) – (المراجع)

يجهلون الإجابة الصحيحة كما الباحث. وثمة طريقة أفضل وتشمل تجربة ضابطة (*) حيث يقوم أحد المعلمين بمقارنة تقدم التلاميذ الذين تعلموا القراءة باستخدام الطريقة الصوتية مع تقدم تلاميذ آخرين يكافئوهم بالقدرات وتعلموا القراءة باستخدام الطريقة البصرية.

ويتعين أن يشمل تصميم البحث كذلك قائمة بالمقاييس أو الأدوات التي سوف تستخدم في جمع البيانات. ويتحتم على الباحثين إيجاد اختبارات ملائمة، وموازين القياس، وأدوات أخرى لازمة لقياس المتغيرات، كما يتعين تقييم ثبات وصحة تلك الإجراءات. والغاية هي اختيار الإجراءات الموضوعية والموثوقة بقدر الإمكان، دون تضحية بالوفاء للمفاهيم التي يفترض أنها تمثلها.

العينة The Sample

يجب أن تشمل الخطة وصفاً للمجتمع الإحصائي ذي العلاقة بالدراسة، وهذا يعني: نوع الأفراد المشمولين بالبحث. ويتحتم على الباحث أن يولي بعض الاهتمام إلى توفر أولئك الأفراد. ومن الضروري أيضاً وصف الإجراءات التي سوف تتبع في اختيار العينة. ويتعين تحديد المجتمع الإحصائي للعينة ، والوسائل التي سوف تتبع في انتقاء العينة وحجمها المقترح.

التحليل الإحصائي The Statistical Analysis

وأخيراً، يتعين على خطة البحث أن توجز الطريقة المقترحة للتحليل الإحصائي للبيانات. وقبل البدء بتجميع البيانات، يجب على المرء أن يحدد الأسلوب الإحصائي الذي سيستخدم لإيجاد جواب لمسألة البحث، أو لاختبار الفرضية. وسوف يكون المرء بحاجة لوصف أو اختصار البيانات المتجمعة من العينة المدروسة. ويتوجب بعدئذ، أن يكون المرء قادراً على تقييم مدى موثوقية ودقة الاستنتاجات والتعميمات من نتائج العينة إلى المجتمع كله. وتخدم الوسائل الإحصائية كلا الوظيفتين. فوظيفة تلخيص البيانات المتوفرة تتم بواسطة الإحصاء الوصفي. أما الإحصاء الاستدلالي فإنه يساعد المرء على الاستدلال من بيانات العينة.

ويجد الكثير من الباحثين المتمرسين، وكذلك أولئك الذين بدأوا، لتوهم، تعلم العملية، ضرورة استشارة أحد الخبراء الإحصائيين قبل إكمال خطة بحثهم. وثمة في الفصل السادس مناقشة وجيزة للدور الذي يلعبه التحليل الإحصائي في اختبار الفرضيات.

(*) التجربة الضابطة (Controlled Experiment): تجربة يجرى بموجبها ضبط المتغيرات المستقلة وتغيير المتغيرات التابعة أثناء سير التجربة. (د. محمد علي الخولي: قاموس التربية) – (المراجع)

الخلاصة SUMMARY

من اجل الشروع في المرحلة التوكيدية للدراسة البحثية، لابد أن تتوفر واحدة أو أكثر من الفرضيات ذات البيان الواضح. فالفرضية هي تنبؤ الباحث حول ناتج الدراسة. وتشتق الفرضيات استقرائياً من الملاحظة، أو استنتاجيا من نظرية معروفة. فالمعرفة والخبرة بالمجال، وحسن الإطلاع على بحوث سابقة، عوامل مهمة في صياغة فرضية مقنعة.

وتخدم الفرضية وظيفة متعددة الأغراض في البحث. ونظراً لأنها تقترح تفسيراً يمكن اختباره تجريبياً، فإنها توسع المعرفة. وتهيئ الفرضية توجيهاً لجهود الباحث إذ إنها تحدد طريقة البحث ونوع البيانات اللازمة ذات الصلة بحل المشكلة. كما إنها تهيئ أيضاً إطاراً لتفسير النتائج، ولبيان استنتاجات الدراسة.

ويتعين على الفرضية الجيدة أن تفي بمعايير معينة: إذ يتوجب أن تكون لها قوة تفسيرية، وأن تكون قابلة للاختبار، وهذا يعني أنها تخص متغيرات يمكن قياسها، ويتوجب على الفرضية أن تتفق مع أكثرية البيانات المتوفرة، ويتحتم أن تكون صياغتها واضحة وموجزة بقدر الإمكان، كما يتوجب عليها بيان العلاقة المتوقعة بين المتغيرات.

وبعد صياغة الفرضية، تكون الخطوة التالية وهي كتابة خطة للبحث التي تشتمل على بيان المشكلة، والفرضية ووصفاً لتصميم البحث، والعينة، والتحليل الإحصائي الذي سوف يطبق. وتتيح الكتابة الابتدائية للخطة فرصة للباحث وللآخرين لتقرير ما إذا كان هناك برنامج عملي لاختبار الفرضية، بالوسع تنفيذه.

وبعد صياغتها وتقييمها باستخدام المعايير المذكورة أعلاه، تكون فرضية البحث مهيأة للاختبار التجريبي. وتتم صياغة الفرضية الصفرية– وهي تتمثل في نفي ما يتوقعه الباحث. ومن المهم أن يتذكر المرء بأنه لا يمكن إثبات أو دحض أية فرضية، فبالإمكان، فقط، دعمها أو عدم دعمها. وحتى إذا لم تلق الفرضية الدعم، فيمكن للفرضية نفسها أن تخدم غاية مفيدة إذ بوسعها أن تقود الباحث إلى إعادة تقييم الأساس المنطقي والإجراءات، والتفكير في أساليب أخرى للمشكلة.

مفاهيم أساسية Key Concepts

criteria for evaluating hypotheses	معايير لتقييم الفرضيات
deductive hypothesis	فرضية استنتاجية
directional hypothesis	فرضية موجهة
functions of hypotheses	وظائف الفرضيات
inductive hypothesis	فرضية استقرائية
nondirectional hypothesis	فرضية غير موجهة

فرضية صفرية	null hypothesis
اختبار الفرضيات	testing hypotheses
نظرية	theory

تمارين EXERCISES

1. ما هي غاية الفرضيات؟
2. ما الفرق بين الفرضية الاستنتاجية والفرضية الاستقرائية؟
3. قم بصياغة فرضية تستند إلى كل من الأسئلة البحثية الآتية:

أ. ماذا سيكون تأثير (طريقة كويزينير) في تعليم الحساب في الابتدائية؟

ب. هل هناك علاقة بين جنس المعلم والتحسين في تحصيل القراءة لدى طلبة سود ذكور في الابتدائية؟

جـ. هل تؤثر المعيشة في سكن مشترك لأعراق مختلفة على اتجاه الفرد تجاه عرق آخر.

د. هل هناك علاقة بين نوعية التعزيز (ملموس أو غير ملموس) ومقدار التعلم الحاصل لدى أطفال محرومين اجتماعياً واقتصادياً؟

هـ. هل يقلل التدريب قبل المدرسة من الفجوة التربوية بين الأطفال المحرومين وغير المحرومين قبل دخولهم الصف الأول الابتدائي؟

و. هل هناك أي تأثير لتوقعات المعلمين بخصوص الأداء العقلي المعرفي للأطفال على واقع أدائهم الحقيقي؟

4. أعد كتابة الفرضية التالية في صيغة صفرية: الأطفال الذين يكون مستواهم في القراءة أقل من مستوى الصف الذي هم فيه سيكون رضاهم عن المدرسة أقل من أولئك الذين مستواهم في القراءة في مستوى (أو أعلى من) الصف الذي هم فيه.
5. قم بتقييم كفاية كل من الفرضيات التالية. وإذ كانت الفرضية غير وافية، فبين سبب عدم كفايتها، وأكتب فرضية كافية.

أ. يستحق المعلمون مرتباً مالياً أعلى من الإداريين.

ب. سيكون الطلبة الذين يدرسون مقرراً حول المدرسة الحكومية المتوسطة، أقدر في إصدار أحكام مستنيرة بخصوص الشؤون السياسية المحلية، من أولئك الطلبة الذين لا يدرسون ذلك المقرر.

جـ. التدريب المستند إلى الحاسوب وكذلك الممارسة، طريقة أفضل لتعليم بطيئي التعليم جدول الضرب من البطاقات الخاطفة.

د. إذا أختلف التلاميذ في خلفيتهم الاجتماعية/ والاقتصادية، فسوف تختلف تقديرات مهاراتهم في اللغة الإنكليزية.

هـ. الأطفال الذين يظهرون دافعية عالية للإنجاز سيكون مستوى قلقهم مرتفعاً حسب مقياس القلق الظاهر عند الأطفال.

و. قيام المعلم بالتعزيز الشفوي الإيجابي لاستجابات التلاميذ، سوف يقلل احتمال الاستجابات المستقبلية.

6. أكتب فرضية موجهة وأخرى غير موجهة تستندان إلى مسألة البحث: ما العلاقة بين معدل نضج مراهقة الأولاد ومفهومهم للذات.

7. لماذا يتوجب صياغة فرضية بوضوح قبل ابتداء البحث؟

8. صنف الفرضيات التالية حسب كونها فرضية بحث أو فرضية صفرية:

أ. سيحصل الطلبة على درجات تحصيل أقل في الاختبارات التي تقيس المستويات الأعلى في تصنيف بلوم.

ب. ليس هناك اختلاف في أداء الطلبة الذين تعلموا الرياضيات بالطريقة (أ) وأولئك الذين تعلموها بالطريقة (ب).

جـ. لا تختلف معدل درجات الحفظ لأطفال يتناولون العقار التجريبي (X) عن معدل درجات الأطفال الذين لا يتناولون ذلك العقار.

د. مهارات حل المشكلة لدى الطلبة الذين يدرسون على أيدي أساتذة متساهلين أعلى من تلك التي لدى طلبة يدرسون على أيدي أساتذة متسلطين.

9. جد دراسة بحثية تبين إحدى الفرضيات، ثم حاول أن تتعرف إلى النظرية التي انبثقت منها تلك الفرضية.

10. أنقد الفرضية التالية: (استخدام طريقة النقاش في الدراسات الاجتماعية بالمدرسة الثانوية سوف ينتج عنه مواطنون كبار أفضل ، من استخدام طريقة المحاضرة).

11. قم بصياغة خطة بحثية تجريبية (مؤقتة) لمشروع صفي:

أ. ما المشكلة العامة للبحث، المطروحة للاستقصاء؟

ب. قم بصياغة المشكلة العامة أعلاه في سؤالي بحثي.

جـ. وضح الأساس المنطقي لدراسة كهذه. ما هي تطبيقاتها النظرية والعلمية؟

د. قم بصياغة الفرضية (أو الفرضيات) لهذه الدراسة.

هـ. هل تم اشتقاق هذه الفرضية استنتاجيا من النظرية، أم استقرائياً من الخبرة والملاحظة؟ وضح إجابتك.

و. قم بتحديد المتغيرات في الدراسة وعرف كل واحد منها إجرائياً.

ز. ما نوع منهجية البحث التي ستحتاجها هذه الدراسة.
ح. ما نوع الأفراد (العينة) الذين ستختارهم للدراسة؟
ط. هل وجدت أي بحث منشور ذي صلة بمشكلتك؟ وإن كان الأمر بالإيجاب، قم بتلخيص نتائج البحث باختصار.

الأجوبة ANSWERS

1. تكمن غاية الفرضيات في توفير مقترح تجريبي (مؤقت) كحل لمشكلة أو كتفسير لبعض الظواهر.

2. يقوم الباحث في حالة الفرضية الاستقرائية بملاحظات حول العلاقات، ثم يفترض بعد ذلك أحد التفسيرات للسلوك الملاحظ. وفي حالة الفرضية الاستنتاجية يقوم الباحث بصياغة إحدى الفرضيات المستندة إلى نظرية معروفة، ويصاحبها في ذلك أساس منطقي للمقترح الخاص.

3 أ. تلاميذ الابتدائية الذين تعلموا بطريقة (كويزينير) سيحصلون على درجات أعلى في اختبار الرياضيات من أولئك الذين لم يتعلموا بتلك الطريقة.
ب. سيحصل تلاميذ الابتدائية من الذكور السود الذين علمهم معلم على درجات أعلى في القراءة من أولئك التلاميذ الذين علمتهم معلمة.
جـ. سيعبر الناس الذي يعيشون في سكنى لأعراق مختلفة عن اتجاهات أكثر إيجابية تجاه أي عرق آخر، من الذين يعيشون في بيوت أحادية العرق.
د. الأطفال المحرمون اجتماعياً/ واقتصادياً الذين يعززون بحوافز ملموسة، سيحصلون على تعلم أكثر من أطفال يجرى تعزيزهم بحوافز غير ملموسة.
هـ. أطفال ما قبل المدرسة المحرومون وغير المحرومين الذين ينالون تدريباً قبل المدرسة ستفصلهم فجوة تعليمية أصغر من أمثالهم الذين لا ينالون تدريباً قبل المدرسة.
و. الأطفال الذين يتوقع مدرسوهم أن يكون أداؤهم العقلي المعرفي عال، سيكون مستوى أدائهم أعلى من الأطفال الذين يتوقع مدرسوهم أن يكون أداؤهم العقلي المعرفي متدن.

4. ليس هناك فرق في الرضا عن المدرسة بين أطفال مستوى قراءتهم أدنى من المستوى المطلوب، وأطفال مستوى قراءتهم حسب المستوى المطلوب أو أعلى منه.

5 أ. إن الفرضية ليست كافية لأنها عبارة قيمية، ولا يمكن استقصاؤها في دراسة بحثية. والفرضية الصحيحة تكون كالآتي:

سوف يعبر المعلمون الذين يتسلمون مرتبات أعلى من الإداريين عن رضا وظيفي أعلى من معلمين لا يتسلمون مرتبات أعلى من الإداريين.

ب. الفرضية ليست كافية لأن (الأحكام المستنيرة) مصطلح قيمي. والفرضية المقبولة تكون كالآتي: سيثبت الطلبة الذين يختارون مقرر المدرسة الحكومية المتوسطة أن لديهم معرفة أكثر في الشئون السياسية المحلية وأنهم يتوصلون في أغلب الحالات إلى استدلالات تقوم على هذه المعرفة أكثر من أولئك الذين لا يختارون مقرر المدرسة الحكومية المتوسطة.

جـ. الفرضية غير وافية لأن كلمة (أحسن) مصطلح قيمي، وينقصها التعريفات الإجرائية الواضحة الموجزة. والفرضية الصالحة للاختبار تكون كالآتي: (الطلبة الذين أداؤهم دون المستوى الدراسي، الذين يتدربون على عمليات الضرب بالحاسوب والممارسة ستكون في المتوسط، نسبة إجاباتهم الصحيحة في اختبار للإتقان، أعلى من طلبة أداؤهم دون المستوى الدراسي ويقضون الوقت ذاته بالتدريب على عمليات الضرب باستخدام البطاقات الخاطفة.

د. الفرضية غير وافية، بسبب عدم وجود بيان لعلاقة متوقعة بين المتغيرات، والفرضية المقبولة تكون كالآتي: (أن الطلبة الذين قد تم تصنيفهم بانتمائهم إلى خلفية اجتماعية/ واقتصادية مرتفعة، سوف يحصلون على درجات أعلى في اختبار المهارة باللغة الإنجليزية، من طلبة آخرين تم تصنيفهم بانتمائهم إلى خلفية اجتماعية/ واقتصادية منخفضة).

هـ. الفرضية غير وافية، إذ لا توجد متغيرات مستقلة أو تابعة. والفرضية المقبولة تكون كالآتي: (الأطفال الذين لديهم دافعية عالية للإنجاز سوف يحصلون على درجات في مقياس (القلق الظاهر عند الأطفال)، أعلى من أطفال ذوي دافعية متدنية للإنجاز).

و. الفرضية غير مناسبة، إذ أنها لا تتماشى مع المعرفة السائدة بقوة التعزيز الإيجابية وتأثيرها على استجابات الطلبة.

6. الفرضية الموجهة: الأولاد ذوي نضج المراهقة المبكرة سوف يظهرون مفاهيم للذات إيجابية أكثر من الصبيان ذوي نضج المراهقة المتأخر.

الفرضية غير الموجهة: هناك اختلاف في مفاهيم الذات للصبيان ذوي نضج المراهقة المبكرة والمتأخرة.

7. تقدم الفرضية، التوجيه لتجميع وتفسير البيانات. فالصياغة الواضحة للفرضية تكشف عيوباً لم تكن ظاهرة أثناء تطوير الفكرة الغامضة للدراسة في المخيلة.

8 أ. بحثية.

ب. صفرية.

جـ. صفرية.

د. بحثية.

9. ستتنوع الإجابات.

10. الفرضية غير صالحة للاختبار خلال فترة معقولة من الوقت. وإضافة إلى ذلك، سوف يكون من الصعوبة تحديد وقياس (مواطنون كبار أفضل).

11. ستتنوع الإجابات.

REFERENCES المصادر

Coopersmith, S. (1981). The antecedents of self-esteem. San Francisco: Consulting Psychology.

Darwin, F. (Ed.). (1898). The life and letters of Charles Darwin (Vol. 1). New York: Appleton.

Davis, F (1988). The effect of time of day of instruction on eighth-grade students' English and mathematics achievement. High School Journal, 71, 78-80.

Hansen, J., and Pearson, P.D. (1983). An instructional study: Improving the inferential comprehension of good and poor fourth-grade readers. Journal of Educational Psychology, 75, 821-829.

Hines, C.V, Cruickshank, D.R., and Kennedy, J.J. (1985). Teacher clarity and its relationship to student achievement and satisfaction. American Educational Research Journal, 22, 87-99.

Piaget, J. (1968). Six psychological studies. New York: Vintage Books.



الجزء الثالث

التحليل الإحصائي

STATISTICAL ANALYSIS

* الإحصاء الوصفي.
* المعاينة والاستدلال الإحصائي.



الإحصاء الوصفي

Descriptive Statistics

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون بوسع الطالب أن:

1- يحدد سمات وحدود أربعة أنماط من المقاييس – الأسمى، الترتيبي، الفتري، والنسبي.

2- ينظم بيانات البحوث في توزيعات تكرارية وعرضها كمضلعات تكرارية ومدرجات تكرارية.

3- يميز بين مقاييس النزعة المركزية والأوضاع التي يجب أن تستخدم فيها وقياس الوسط الحسابي والوسيط والمنوال.

4- يصف التطبيقات المناسبة لقياسات التشتت/ الانتشار وحساب التباين والانحراف المعياري والانحراف الربيعي ومدى أية مجموعة من البيانات.

5- يحوّل الدرجات الخام إلى درجات معيارية ويحدد الموقع النسبي للدرجات المعيارية المقننة (Z) في المنحنى الاعتيادي/ الاعتدالي.

6- يحدّد التطبيقات المناسبة لمؤشرات/ معاملات الارتباط المتنوعة لوصف العلاقة بين المتغيرات.

7- يحسب معامل الارتباط لمجموعتين من البيانات الفترية والنسبية.

8- يصف معنى معامل التحديد/ القرار ويطبقه في تفسير معامل الارتباط.

9- يحدّد معنى وخصائص وتطبيقات تحليل التحليل (meta-analysis).

الإجراءات الإحصائية هي في الأساس طرق لمعالجة المعلومات الكمية بطريقة تقيم معنى لتلك المعلومات. ولهذه الإجراءات ميزتان رئيستان. أولاً إنها تمكننا من تنظيم وتلخيص ووصف ملاحظاتنا، وتدعى مثل هذه الأساليب بالإحصاء الوصفي. ثانياً إنها تساعدنا على تحديد مدى الثقة التي يمكننا فيها استنتاج أن تلك الظواهر التي لوحظت في مجموعة محدودة، "عينة"، ستحدث كذلك في مجتمع أكبر لا نلاحظه، وهو موضع اهتمامنا، الذي نختار العينة منه - بعبارة أخرى، ما مدى الدقة التي يمكننا فيها استخدام التفكير الاستقرائي لاستنتاج أن ما نلاحظه في الجزء سيُلاحظ في الكل. فمع مشكلات من هذا النوع، سوف نحتاج إلى استخدام الإحصاء الاستدلالي.

وتعتبر معرفة بعض الإجراءات الإحصائية الأساسية جوهرية لأولئك الذين يعتزمون تنفيذ البحوث ليتمكنوا من تحليل وتفسير بياناتها وإيصال نتائجهم إلى الآخرين. إضافة إلى ذلك، فإن من المحبذ أن يكون التربويون، الذين يحتاجون إلى مواكبة البحوث والاستفادة من نتائجها، على بينة بالإجراءات الإحصائية كي يتمكنوا من فهم وتقييم دراسات البحوث التي يجريها الآخرون. فالتطبيق الصحيح والتفسير للاختبارات المستخدمة في مدارسنا، تتطلب بعض الفهم للإجراءات الإحصائية. فالمدرسون غير الملمين بهذه الإجراءات قد يجدون صعوبة في تقييم قدرات وتحصيل تلاميذهم. ويجدون من الصعب، كذلك، مراجعة البحوث في ميادين اختصاصهم والحصول على معلومات حديثة.

المقاييس العلمية / سلالم القياس
SCALES OF MEASUREMENT

إن الخطوة الأساسية في إجراء البحوث هي القياس: وهو العملية التي تترجم فيها الملاحظات إلى أعداد. لقد أوضح (S. Stevens, 1951.S) بأن "القياس، في أوسع معنى له، هو تخصيص الأعداد للأشياء أو الأحداث طبقاً للقواعد". فالباحثون يبدأون بالمتغيرات ثم يستخدمون القواعد لتحديد كيفية التعبير عن هذه المتغيرات بشكل عددي. فمتغير الأفضلية الدينية قد يقاس طبقاً للأرقام الموضحة من الطلبة الذين يطلب منهم اختيار واحدة من بين الأفضليات التالية: 1) كاثوليكية، 2) يهودية، 3) بروتوستانتية، 4) أخرى. أما متغير الوزن، فقد يقاس بأعداد تتم ملاحظتها عندما يصعد الأفراد على الميزان. وقد يقاس متغير النضج الاجتماعي كدرجات على مقياس فاينلاند للنضج الاجتماعي.

إن طبيعة عملية القياس التي تبرز الأعداد، تحدد التفسير الذي ينشأ عنها والإجراءات الإحصائية التي يمكن استخدامها معها بشكل مفيد. وأفضل تصنيف مقتبس بشكل واسع من إجراءات القياس هو مقاييس / سلالم ستيفنس Stevens للقياس التي يصنف فيها القياس كاسمي وترتيبي وفتراتي ونسبي.

المقياس الاسمي Nominal Scale

إن أبسط مقياس علمي للقياس هو المقياس الاسمي (أو النوعي)، وهو يتضمن تصنيف الأشياء أو الأفراد في أصناف مختلفة نوعاً وليس كماً. ويتطلب القياس في هذا المستوى أن يكون المرء في نهاية الأمر قادراً على تحديد أصناف عناصر الدراسة بكونها، منفصلة بالتبادل(*)، وقادراً على تحديد المعايير التي تعين الصنف المناسب للعنصر. إن العملية التجريبية المطلوبة على هذا المستوى من القياس تنطوي فقط على إدراك ما إذا كان الشيء أو الفرد ينتمي أو لا ينتمي إلى صنف معين. إن العلاقة الوحيدة بين الأصناف هي أنها مختلفة بعضها عن الآخر، وليس هناك أي إيحاء بأنها تمثل "أكثر" أو "أقل" من السمات. فتصنيف الطلبة وفق الجنس (ذكر أو أنثى)، يشكل قياساً أسميا/ نوعياً.

قد تستخدم الأعداد على مستوى اسمي ولكن عند اختيارها، فقط، تعتبر رموزا لتمييز الأصناف، وليس هناك علاقة تجريبية بين الأعداد المستخدمة في القياس الاسمي بما ينطبق مع العلاقة الرياضية الفعلية بين الأعداد المستخدمة. فمثلاً، قد يستخدم الرقم (صفر) ليمثل الذكور والرقم (1) ليمثل الإناث – إلا أن (1) لا يوضح أكثر مما يوضحه (صفر) وإن الأرقام يمكن تبادلها دون أي تأثير إلا على نظام التصنيف المستخدم. فالمنطقة التعليمية (231) هي ليست بالضرورة أكثر أو أقل من المنطقة التعليمية (103). فالأعداد المستخدمة على المستوى الاسمي لا تمثل أي مقدار مطلق أو نسبي لأي سمة بل مجرد رموز تحدد عناصر صنف معين. إن الأعداد المحددة في المقياس الاسمي لا يمكن، بطبيعة الحال، استخدامها رياضياً من خلال الجمع أو الطرح أو الضرب أو القسمة. فقد يستخدم المرء تلك الإجراءات الإحصائية فقط، استنادا على مجرد العد، كذكر أعداد الملاحظات في كل صنف أو التعبير عن تلك الأعداد كنسب مئوية للعدد الإجمالي للأفراد.

المقياس الترتيبي Ordinal Scale

الشيء التالي في هرم معايير القياس هو المقياس الترتيبي الذي يصنف الأشياء أو الأشخاص وفق مدى كثرة / قلة الصفة الخاصة بهم التي هي موضع الدراسة. فالأشياء والأشخاص يوضعون في ترتيب وفق درجة الصفة الخاصة بهم – لكن، دون إيضاح درجة الفرق أو المسافة بينهم، فمثلاً يستخدم المقياس الترتيبي حين يرتب المدرسون تلاميذهم من الأعلى إلى الأدنى حسب خصائص مثل النضج الاجتماعي أو قدرات القيادة أو التعاون.

فالحاجة الأساسية للقياس في هذا المستوى هو وجوب أن تكون العلاقة على النحو: إذا كان الشيء (أ) أكبر من الشيء (ب) وأن الشيء (ب) هو أكبر من الشيء (جـ) فإن الشيء (أ)

(*) الانفصال بالتبادل (mutually conclusive): مصطلح يشير إلى عدم وجود أي عنصر مشترك بين صنفين/ مجموعتين – (المراجع).

أكبر من الشيء (جـ). وهي تكتب كما يلي:

إذا كان (أ > ب) و (ب > جـ) فعندئذ (أ > جـ). وبشكل مناسب، قد تستخدم صيغة لتحل بدلاً من (أكبر من) مثل – أقوى من، يسبق، له أكثر من، وهكذا.

إن الإجراء التجريبي في القياس الترتيبي يتضمن المقارنة المباشرة للأشياء أو الأفراد بدلالة مدى الميزة المعنية وحدها التي تعود لهم. وهكذا، عندما يتم تحديد الأعداد في القياس الترتيبي فأنها توضح، فقط، "ترتيب" الملاحظات ولا أكثر من ذلك. فلا يوجد معنى للفرق بين الأعداد ولا النسبة بينها. وهكذا إذا استخدمت الأعداد 1، 2، 3 لتصنيف الملاحظات، فليس هناك أي تضمين بأن المسافة بين الرتب متساوية – أي أن (1) هو أعلى مرتبة من (2) بمقدار ما يكون (2) أعلى من (3). ففي سباق المشي غير المحدد زمنيا، نعرف فقط من يأتي أولاً وثانياً وثالثاً. ولا يوجد سبل لمعرفة كم كان أحد العدائين أسرع من أي عداء آخر. فالفرق بين المركز الأول والثاني لن يكون بالضرورة مماثلاً للفرق بين الثاني والثالث، وليس بوسع المرء أن يقول إن العداء الذي جاء بالمركز الثاني كان أسرع مرتين من العداء الذي جاء بالمركز الرابع. كذلك، فإن ترتيب الطلبة على أساس درجاتهم في الرياضيات أو ترتيب طلبة سنة التخرج على أساس تقديراتهم يمثل المقياس الترتيبي.

ومثال مفيد آخر للمقياس الترتيبي هو مقياس صلابة/ قسوة المعادن. فالمعادن ترتب وفق قدرتها على خدش بعضها. فإذا كان المعدن (أ) قادر على خدش المعدن (ب)، فيقال إن المعدن (أ) أصلب من المعدن (ب)، وعلى هذا الأساس يرتب الماس على أنه الأصلب لأن بوسعه أن يخدش جميع أنواع المعادن المعروفة، ولكن لا يمكن خدشه بالمعادن الأخرى. إن مجموعة من عشرة معادن مرتبة حسب الصلابة من الألين إلى الأصلب، قد انتخبت كمعيار وخصصت أرقام من 10-1 حيث يوضح (1) المعدن الأطرى و (10) المعدن الأصلب. وأعطيت معادن أخرى أعداداً على أساس اختبار الخدش، وهكذا تعرف ترتيب صلابة المعدن لكننا لا نعرف مقدار صلابة معدن عن آخر. ولا يمكن الافتراض أن المعدن ذا القيمة (4) هو أصلب مرتين من المعدن ذي القيمة (2) أو أن الفرق في الصلابة بين المعدنين (2) و (4) هو ذات الفرق في الصلابة بين المعدنين (1) و (3).

الإحصائيات/البيانات الإحصائية المناسبة للمقياس الترتيبي محدودة. ونظراً لأن مقدار الفترة بين الأصناف غير معرف، فلا يمكن استخدام أي إجراء إحصائي يفترض فترات متساوية. فالإحصائيات، التي توضح النقاط التي تحتها نسب مئوية معينة للحالات، مناسبة للمقياس الترتيبي.

مقياس الفترات Interval Scale

هناك نظام قياس أكثر دقة وهو مقياس الفترات الذي لا يرتب الأشياء أو الأحداث وفق مقدار

الميزات/ الصفات التي تمثلها، فحسب، بل كذلك له أصل (صفر) كيفي ويقيم فترات متساوية بين وحدات القياس. فالفروق المتساوية في الأعداد تمثل فروقاً متساوية في الميزة المقاسة. ويعتبر مقياس الحرارة الفهرنهايتي والمئوي أمثلة على مقاييس الفترات. ففي مقياس الفترات، هناك معنى للترتيب ولعلاقات المسافة بين الأعداد. فيمكننا التأكيد أن الفرق بين الدرجتين المئويتين (50) و (51) مساوٍ للفرق بين الدرجتين المئويتين (30) و (31). ولا يسعنا القول، على أية حال، أن حرارة الدرجة (50) تعادل ضعف حرارة الدرجة (25). ويعزى ذلك إلى أنه لا توجد "نقطة صفر حقيقية" في ميزان الفترات. فنقطة الصفر تقوم على الاتفاق،كما في الميزان المئوي، الذي يحدد القيمة صفر لدرجة تجمد الماء. وبشكل مماثل فإن نقطة الصفر على مقياس اختبار نفسي أو تربوي هي كيفية. مثلاً، ليس هناك ذكاء صفري، وليس هناك طريقة في اختبارات الذكاء المقننة لتحديد شخص بذكاء صفري. فالتلميذ قد يحصل أحياناً على الدرجة صفر في اختبار الإحصاء لكن ذلك لا يعني أن معرفته "صفر" في الإحصاء. ولو كان لدينا ثلاثة تلاميذ حصلوا على معدلات (15) و (30) و (45) في اختبار الإحصاء فلا يسعنا القول أن الدرجة (30) تمثل معرفة إحصائية بمقدار مرتين للدرجة (15) أو إن الدرجة (45) تمثل معرفة هي ثلاثة أضعاف الدرجة (15). ولفهم سبب ذلك، دعنا نفترض أن (15) بنداً بسيطاً جداً تضاف إلى الاختبار بحيث إن جميع التلاميذ الثلاثة يستطيعون الإجابة عليها بشكل صحيح، وستصبح الدرجات الثلاث الآن (30) و (45) و (60) للطلبة الثلاثة. ولو حاولنا صياغة نسب بين القيم على مقياس الفترات هذا، فأننا سنذكر (بشكل خاطئ) إن التلميذ الذي درجته (60) له معرفة بالإحصاء ضعف التلميذ الذي درجته (30) لكن في النسبة السابقة كنا قد افترضنا (بشكل غير صحيح) أن للتلميذ معرفة بالإحصاء تعادل ثلاثة أضعاف ما لدى التلميذ الآخر، وهكذا لا تتفق النتائج.

بالإمكان إجراء عمليتي الجمع والطرح على أعداد مقياس الفترات، ونظراً لأن الصفر "كيفي" فإن عمليتي ضرب وقسمة الأعداد غير مناسبتين، فكما رأينا فإن النسب بين الأعداد على مقياس الفترات تخلو من المعنى. وعلى أية حال، أن الفرق بين المواقع على مقياس الفترات قد تذكر أو إن الأعداد قد تجمع. وقد تستخدم الإجراءات الإحصائية القائمة على أساس الجمع مع مستوى المقياس هذا، إضافة إلى الإجراءات المناسبة لمقاييس المستوى الأدنى. وتشمل هذه أغلب الإجراءات الإحصائية العامة.

من المهم الإشارة إلى أنه في أغلب المقاييس الأكاديمية، تكون الفترات متساوية في إطار أداة القياس ذاتها، ولكن ليس بالضرورة في إطار القدرة التي نقيسها. ولإيضاح ذلك، لندرس اختبار الهجاء مع الكلمات التالية: قطة، صحن، كرة، ثرثار، شيزوفرينيا، فقر الدم. وهنا تكون المسافة بين (1) و (3) الصحيحين هي نفس المسافة بين (3) و(5) الصحيحين. وعلى أية حال، إذا درسنا الفرق في إطار قدرة التهجي بين (3) و (5) الصحيحين فإنه يشير إلى فرق أكبر في

القدرة من الفرق بين (1) و (3) الصحيحين. فما لم يسعنا القول بأن المسافة بين (3) و (5) في اختبار التهجي يمثل نفس المقدار من قدرة التهجي كالمسافة بين (1) و (3) فإن هذه النقاط لا توضح شيئاً أكثر من ترتيب تدريجي التلاميذ.

لكنه، بالإمكان من خلال البناء الدقيق إنتاج أداة حيث تعطى الفترات بين درجات الاختبار تقريباً معقولاً لفترات القدرة. وأجود اختبارات الذكاء، مثال على ذلك. فالفرق في القدرة بين معامل الذكاء (90) ومعامل الذكاء (95) قد لا يكون ذات الفرق بين معامل الذكاء (105) ومعامل الذكاء (110) إلا أننا قد لا نخطئ كثيراً إذا افترضنا أن الفرقين متساويان تقريباً.

لقد أصبح الأمر ممارسة مألوفة إذ تجري معالجة العديد من المتغيرات التربوية كالاختبارات الصفية والتقديرات (أ = 4، ب = 3، وهكذا) كما لو كانت بيانات فترات حتى وإن كان الافتراض غير مبرر بشكل جيد. فمن الصعب تأكيد أن الفرق بين هـ = صفر و د = 1 يمثل ذات الفرق في التحصيل الأكاديمي كالفرق بين جـ = 2 و ب = 3 أو تسويغ معالجة النقاط في مثالنا عن اختبار الهجاء كفترات. فلابد أن يكون المرء حذرا عند تفسير إحصائيات الفترات المستمدة من هذه البيانات. فالإحصائيات تنطوي على معلومات بمستوى الفترات عندما تكون هذه المعلومات في مكان ما بين الترتيب والفترات.

المقياس النسبي Ratio Scale

إن المقياس النسبي، هو المستوى الأعلى لمعايير القياس، إذ له نقطة "صفر حقيقي" وفترات متساوية. ويمكن تشكيل النسب بين أي قيمتين معينتين على المقياس. فالمقياس المعياري المستخدم لقياس الطول في وحدات الأمتار أو السنتيمترات هو مقياس نسبي لأن الأصل في القياس هو الصفر المطلق الذي يدل على انعدام الطول في حالات خاصة(*). وهكذا فمن الممكن الإيضاح بأن عصا طولها 60 سم هي أطول مرتين من عصا طولها 30 سم. ومع المقياس النسبي فإنه من الممكن، ضرب أو قسمة كل من القيم بأي رقم معين دون تغيير خصائص المقياس.
- فمثلاً يمكننا أن نضرب (2) متر في (100) كي تغير وحدة القياس إلى (200) سم، أو ضرب (6) أمتار في (10) لنغير الوحدة إلى ديسمترات، ويمكن أن نضرب ونحافظ على النسبة نفسها كالسابق قبل الضرب، مثلاً، يمكن أن نضرب (4) كوارت (ربع جالون) من الحليب و (2) كوارت من الحليب في (2) ونغير وحدة القياس إلى باينت (نصف كوارت)، وبالباينتات فإن كل (8) باينت لا تزال تعادل مرتين الـ (4) باينت.

(*) الصفر المطلق (Absolute Zero): مصطلح يشير إلى صفر مختلف عن "الصفر الكيفي" الخاص بمقياس الفترات، وفي هذه الحالة يكون هناك معنى لانعدام الميزة/ الصفة، فالمسافة بين نقطتين منطبقتين تساوي الصفر، فهناك معنى لانعدام المسافة أو انعدام الطول بين نقطتين منطبقتين (المراجع).

وهناك متغيرات قليلة في مجال التعليم تكون نسبية في طبيعتها. وهي محصورة بشكل واسع بالأداء الحركي والمقاييس الفزيولوجية الأخرى. وبينما يمكن أن نقول أن الشخص بطول (6) أقدام هو مرتين أطول من شخص بطول (3) أقدام لأن الأطوال هي بيانات نسبية، ولا يسعنا أن نقول أن الشخص الذي ينال (40) درجة في اختبار الحساب هو أجود في الحساب مرتين من شخص ينال (20) درجة لأن درجات اختبار الحساب هي بيانات فترات.

إن جميع أنواع الإجراءات الإحصائية مناسبة للمقياس النسبي.

تنظيم بيانات البحث ORGANIZING RESEARCH DATA

إن البيانات الوصفية التي لم يجر تنظيمها حسب نوع من الترتيب، تكون بالغة الصعوبة، إن لم تكن مستحيلة. وعليه، فإن تنظيم البيانات البحثية هو خطوة أساسية في الإحصاء الوصفي. وهناك طريقتان مستخدمتان في الغالب في تنظيم البيانات وهما : 1) ترتيب القياسات في توزيعات تكرارية و 2) عرضها في شكل بياني.

التوزيعات التكرارية Frequency Distributions

إن الترتيب المنتظم للقياسات الإفرادية من الأدنى إلى الأعلى يدعى بالتوزيع التكراري. وينطوي استخدام هذا الأسلوب على مجرد صنع قائمة بالقياسات الإفرادية في عمود مع أعلى قياس في القمة ثم يليه الثاني في العلو من القمة واستمرار في النزول نحو القياس الأدنى الذي يسجل في أسفل العمود. وقد وجد في الغالب أن العديد من الدرجات المتطابقة سوف تقع في توزيع معين. وبدلاً من إدراج هذه الدرجات بشكل منفصل، فإنه من المعتاد إضافة عمود ثان حيث يسجل تكرار كل قياس. ففي الجدول 5.1 هناك درجات مجموعة من (105) طلاب في اختبار الإحصاء. في القسم (أ) من الجدول تدرج الدرجات في شكل غير منتظم. وفي القسم (ب) ترتب نفس الدرجات في توزيع تكراري.

ويمكن من أي توزيع تكراري أن نختبر "الشكل" العام للتوزيع. ومع الدرجات المرتبة بهذه الصورة يمكن للمرء أن يحدد انتشارها، أي ما إذا كانت موزعة بانتظام أو تميل إلى التجمع وتحديد مواقع التجمعات في التوزيع. فمثلاً عند النظر إلى التوزيع التكراري للدرجات المعروضة في الجدول 5.1، فإنه من السهل أن نرى أنها تتراوح من (21) إلى (35) وإن (29) هي الدرجة الأكثر تكراراً، وان هناك نزوعا للدرجات في التجمع أكثر قرب قمة التوزيع مما هو عند القاعدة. ولن يتضح أي من هذه لو كانت الدرجات منتظمة. فتنظيم البيانات في توزيعات تكرارية يسهل كذلك من حساب إحصائيات مفيدة متنوعة.

جدول (5.1): درجات اختبار 105 طلاب في اختبار الإحصاء

أ- درجات غير منظمة

23	29	27	31	34	24	28	32	33	29	23	30	30	29	33
31	32	32	29	26	28	30	26	33	33	26	27	24	29	25
29	27	31	30	35	32	30	31	29	21	33	31	30	34	28
22	33	34	31	30	27	28	30	28	28	22	32	33	29	26
31	29	31	21	28	33	31	31	21	30	24	29	27	29	30
30	30	33	22	30	29	28	31	32	32	31	29	22	33	31
33	27	28	29	30	21	27	33	27	29	28	33	31	33	32

ب- توزيع تكراري

الدرجات (X)	عدد التكرارات	التكرارات (f)
35	/	1
34	///	3
33	𝍸 𝍸 𝍸	15
32	𝍸 ///	8
31	𝍸 𝍸 ////	14
30	𝍸 𝍸 ////	14
29	𝍸 𝍸 𝍸 /	16
28	𝍸 𝍸	10
27	𝍸 ///	8
26	////	4
25	/	1
24	///	3
23	/	1
22	///	3
21	////	4
		105 = N

عروض بيانية Graphic Presentations

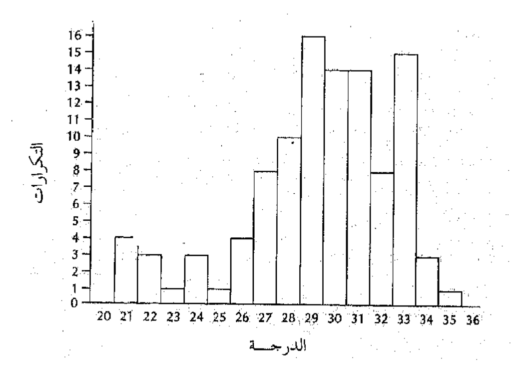
من المناسب غالباً عرض البيانات البحثية في شكل بياني. والأكثر استخداماً بين الأنماط المختلفة للرسوم البيانية هما المدرج التكراري والمضلع التكراري. والخطوات الأولية في بناء المدرج التكراري والمضلع التكراري متطابقة:

1- رتب نقاط الدرجات على المحور الأفقي من أدنى قيمة على اليسار إلى الأعلى في اليمين. أترك مجالاً كافٍ لدرجات إضافية في كلا النهايتين للتوزيع.
2- رتب تكرارات الدرجة (أو الفترات) على المحور العمودي، والترقيم إلى أعلى بدءاً من الصفر.
3- ضع نقطة فوق مركز كل درجة على مستوى تكرار تلك النقطة.

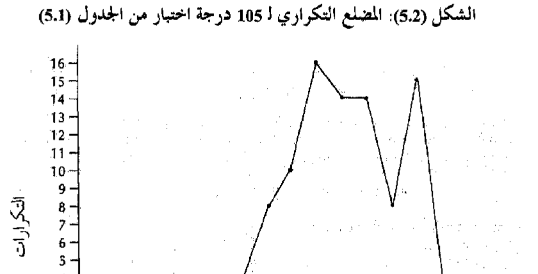
ومن هذه النقطة، يمكن للمرء أن ينشي مدرجاً تكرارياً أو مضلعاً تكراريا. ففي إنشاء المدرج التكراري، ينبغي على المرء أن يرسم خلال كل نقطة خطاً أفقياً مساوٍ للعرض الممثل للدرجة كما هو مبين في الشكل 5.1. ولإنشاء المضلع التكراري، ينبغي وصل النقاط المتجاورة، كما يتم وصل النهايتين للشكل الناتج بالقاعدة (الخط الصفري) مع النقاط التي تمثل الدرجة التي تقل بدرجة واحدة عن أدنى درجة، والدرجة التي تزيد بدرجة واحدة عن أعلى درجة، كما هو مبين في الشكل 2-5.

والمدرجات التكرارية مفضلة حين يود المرء إيضاح الطبيعة المنفصلة للبيانات مثلماً يتم لدى استخدام المقياس الاسمي. وتفضل المضلعات عندما يود المرء إيضاح الطبيعة المستمرة (المتصلة) للبيانات.

الشكل (5.1): مدرج تكراري لـ 105 درجات اختبار في الجدول (5.1)



الشكل (5.2): المضلع التكراري لـ 105 درجة اختبار من الجدول (5.1)



مقاييس النـــزعة المركزية
MEASURES OF CENTRAL TENDENCY

إن إحدى الطرق المناسبة لتلخيص البيانات هو إيجاد مؤشر واحد يمكن أن يمثل مجموعة كاملة من المقاييس، فمثلاً إيجاد درجة واحدة بوسعها أن تشير إلى أداء مجموعة من 300 طالباً في اختيار الاستعداد سيكون مفيداً لأغراض المقارنة. وفي الإحصاء تتوفر ثلاثة مؤشرات لمثل هذا الاستخدام. وهي تدعى "مقاييس النزعة المركزية" أو "المتوسطات". ولأغلب الناس يعني مصطلح "المتوسط" كمجموع الدرجات مقسوماً على عدد الدرجات. أما بالنسبة للإحصائي فإن المتوسط يمكن أن يكون هذا المقياس، المعروف بالمتوسط / الوسط الحسابي، أو أحد المقياسين الآخرين للنزعة المركزية والمعروفين "بالمنوال" أو "الوسيط". ويفيد كل من هذه الثلاثة كمؤشر لتمثيل المجموعة ككل.

The Mode المنوال

المنوال هو القيمة الأكثر تكراراً في التوزيع ، وهي الأبسط من حيث إيجادها من بين المقاييس الثلاثة للنزعة المركزية لأنها تتحدد بطريق التفتيش وليس الحساب. وفي توزيع الدرجات التالي:

14 16 16 17 18 19 19 19 21 22

يمكن لأحدهم أن يرى أن المنوال لهذا التوزيع هو (19) لأنها الدرجة الأكثر تكراراً. وأحياناً هناك أكثر من منوال في التوزيع، فمثلاً، إذا كانت الدرجات:

14 16 16 16 18 19 19 19 21 22

فإن لدينا منوالين 16، 19. إن هذا النوع من التوزيع ذي النمطين يدعى (ثنائي المنوال). والتوزيعات ذو الثلاثة منوالات(*) أو أكثر تدعى (ثلاثية المنوال) أو (متعددة المنوال) على التعاقب.

ويعد المنوال، المؤشر الأقل فائدة للقيمة المركزية في التوزيع لسببين. أولاً: أنه غير مستقر. فمثلاً، قد يكون لعينتين عشوائيتين مستمدتين من ذات المجموعة، منوالان مختلفان تماماً. ثانياً قد يكون للتوزيع أكثر من منوال. ونادراً ما يذكر في البحوث المنشورة المنوال كمؤشر للنزعة المركزية. ففائدته مقصورة بشكل كبير على الأغراض الاستقصائية. ويمكن التعرض للمنوال في أي من معايير القياس، إلا أنها المقياس الوحيد للنزعة المركزية التي قد تستخدم بشكل صحيح في المقياس الأسمي.

الوسيط The Median

يعرف الوسيط على أنه النقطة في توزيع القياسات التي تقع دونها 50 في المئة من الحالات (وهذا يعني أن الـ 50 في المئة الأخرى ستكون فوق هذه النقطة). فمثلاً، إذا أخذنا توزيع الدرجات التالية:

14 16 16 17 18 19 19 19 21 22

فإن النقطة التي تقع دونها 50 في المئة من الحالات هي في الوسط ما بين (18)، (19) وهكذا، فإن الوسيط لهذا التوزيع هو (18.5). ولإيجاد هذه القيمة وضعنا أولاً درجات التوزيع العشر في نسق ترتيبي (أي من الأدنى إلى الأعلى) ثم وجدنا النقطة التي يقع دونها نصف الدرجات. وهذه النقطة (18.5) التي تفصل بالضبط القيمتين (18) و (19) تدعى بالاصطلاح الإحصائي (الحد الأعلى) للدرجة (18) و (الحد الأدنى) للدرجة (19). ولحساب الوسيط، نعتبر أن كل درجة تمثل مدى أو فترة من منتصف المسافة بين تلك الدرجة والدرجة الأدنى التالية. وهكذا، ففي المثال، من المعتقد أن (18) تمثل الفترة من (17.5) حتى (18.5) بينما تمثل (19)، الفترة من (18.5) حتى (19.5) .

من المهم أن نلاحظ أن الوسيط لا يقع دائماً على الحد الفاصل بين القيمتين. فهو، في الحقيقة، يقع في الغالب في موضع ما، بين الحد الأعلى والحد الأدنى للفترة. وبهدف تحديد الوسيط، يجب أن نتصور أية درجة مسجلة، بكونها تمثل المدى بين حديها الأدنى والأعلى وليس نقطة واحدة. أنظر المثال التالي:

(*) منوال (Mode): هناك من يجمعها "مناويل" وهناك من يجمعها على "منوالات" (المراجع).

23 24 25 26 26 26 26 27

في هذا التوزيع يقع الوسيط بين أثنتين من الدرجات (26). لاحظ تكرار الدرجة (26). هناك أربع من هذه الدرجات في التوزيع، واحدة منها تقع دون نقطة الوسط وثلاثة فوقها. وفي هذه الحالة، يجب لإيجاد الوسيط أن نقسم الفترة 25.50 – 26.50 إلى أربعة أقسام، والمسافة بين كل من هذه الدرجات ستكون 0.25 من الفترة. ونعتبر أن كلاً من هذه الدرجات يمثل مدى يغطي 0.25 من المسافة بين 25.50 و 26.50 اللذين هما الحدان الأدنى والأعلى للدرجة 26. ونوضح هذا المفهوم بعرض الموضع والقيمة للمنوال.

			0.25		0.25		0.25		0.25		
23	24	25 ↓	26	↓	26	↓	26	↓	26	↓	27
		25.50		25.75		26.00		26.25		26.5	
		↑ الحد الأدنى		↑ الوسيط						↑ الحد الأعلى	

وبدلاً من المرور بهذه العملية كل مرة، فإن من السهل إيجاد الوسيط باستخدام الصيغة/ القانون كما يأتي:

$$Md = L + \left(\frac{\frac{N}{2} - cfb}{fw} \right) i \qquad (5.1)$$

حيث:

Md = الوسيط

L = الحد الأدنى للفترة التي يقع الوسيط فيها

N = عدد الحالات في التوزيع

cfb = التكرار التراكمي في جميع الفترات تحت الفترة التي تحتوي على الوسيط

fw = تكرار الحالات ضمن الفترة الحاوية على الوسيط

i = حجم/ طول الفترة

ففي المثال السابق يعتبر L= (25.50) و N= (8) و cfb= (3) و fw= (4) و i هو (1) وهكذا.

$$\frac{23 \quad 24 \quad 25}{cfd\,3} \downarrow \frac{26 \quad 26 \quad 26 \quad 26}{fw = 4} \quad \frac{27}{0}$$

$$L = 2550$$

وبوضع هذه القيم في المعادلة نحصل على ما يلي:

$$Md = 25.50 + \left(\frac{\frac{8}{2} - 3}{4} \right) 1 = 25.75$$

وتصبح القيمة (1) ضرورية فقط حين يكون طول الفترة مختلفاً عن (1). فمثلاً، إذا كنا نتعامل مع درجات اختبار تحصيل لمستوى دراسي التي يعبر عنها بالوحدات العشرية مثل (3.4) و (5.9) فإن حجم الفاصلة هو (0.1).

ويمكن استخدام التوزيع التكراري بكفاءة لحساب الوسيط. ويبين الجدول (5.2) نفس البيانات نفسها كالجدول (5.1) ويشمل عموداً إضافياً يدعى ؛ (التكرارات المتراكمة) (cf) المستخدم لتحديد موقع الوسيط. وتبين التكرارات المتراكمة تكرارات القيم صعوداً لتشمل أية فترة معينة في التوزيع. مثلاً، تكرار الدرجات صعوداً وبشمول فترة الدرجة (30) في الجدول (5.2) هي (64) كما مبينة في العمود *cf*. ويخولنا هذا العمود تحديد موقع الفترة التي يقع الوسيط ضمنها. ولأن الوسيط هو تلك النقطة في التوزيع التي يقع 50 في المئة من الحالات تحتها، فإن الفترة التي تحتوي على الوسيط ستكون ذات التكرار التراكمي الذي يحتوي القيمة (N/2). بعبارة أخرى، نقسم N في التوزيع على (2) ونبحث عن الفترة التي تحتوي على هذه القيمة.

وفي توزيع النقاط المبينة في الجدول (5.2) فإن عدد الحالات أو N يساوي (105) وهكذا فأن قيمة N/2 هي (52.50). والنظر في العمود *cf* نرى إن هناك (50) حالة صعوداً وتشمل النقطة (29) و (64) حالة صعوداً نشمل النقطة (30). وهكذا فالمتوسط يقع ضمن الفترة الممثلة بالدرجة (30)، ومن المؤكد الآن تطبيق المعادلة/ القانون (5.1) لإيجاد متوسط التوزيع. إن قيمة L في هذا المثال هي (29.50) لأن الفترة التي تحتوي على الوسيط تمثل المدى من (29.50) كحد أدنى إلى (30.50) كحد أعلى. إن قيمة *cfb* (أي التكرار التراكمي للقيم دون الفترة التي تحتوي على المتوسط) هي (50). إن تكرار القيم ضمن الفاصلة الممثلة بالدرجة (30) (أي قيمة *fw*) هي (14). أما طول الفترة، أو القيمة *i* في التوزيع فهو (1) لأن كل درجة تمثل عرض فترة طولها (1). وبتطبيق المعادلة (5.1) نجد أن متوسط التوزيع هو (29.678).

الجدول (5.2): حساب المتوسط مع توزيع تكرار درجات 105 طلاب في اختبار الإحصاء

الدرجات (X)	التكرارات (f)	التكرارات المتراكمة (cf)
35	1	105
34	3	104
33	15	101
32	8	86
31	14	78
30	fw 14	64
29	16	50
28	10	34
27	8	24
26	4	16
25	1 (cfb = 50)	12
24	3	11
23	1	8
22	3	7
21	4	4

$$\text{(الوسيط)} = Md = 29.50 + \left(\frac{\frac{105}{2} - 50}{14} \right) 1 = 29.678$$

لاحظ أن الوسيط لا يأخذ بنظر الاعتبار حجم الدرجات الفردية. ولغرض إيجاده نرتب البيانات في نسق ترتيبي ونجد النقطة التي تقسم التوزيع إلى نصفين متساويين. والوسيط إحصاءة(*) ترتيبية، لأنه يستند على الرتبة. ويمكن حساب الوسيط من بيانات الفترة أو النسبة، لكن لا تستخدم في هذه الحالات خاصة الفترة للبيانات.

والظرف الوحيد الذي قد يكون فيه الوسيط، القياس المفضل للنزعة المركزية، هو حالة وجود بعض الدرجات المتطرفة/ الحدية في التوزيع. ففي هذه الحالة يؤدي استخدام قياس النزعة المركزية الذي يأخذ بنظر الاعتبار مقدار كل درجة إما إلى مبالغة أو خفض في التقييم للدرجة النموذجية. ونظراً لعدم حساسية الوسيط للدرجات الحدية فإنه يعد مؤشراً مناسباً للتطبيق حين يود المرء إيجاد الدرجة النموذجية. وللإيضاح، أدرس التوزيع التالي:

49 50 51 53 54 55 86 70 89

(*) إحصاءة (Statistic): مصطلح يشير إلى مفردة إحصائية تحدد خاصة للعينة (المراجع).

الدرجة (54) التي هي وسيط التوزيع تكون أفضل درجة نموذجية. والمؤشر الذي يأخذ بنظر الاعتبار القيم الفردية للدرجات (70) و (89) سوف يؤدي إلى مبالغة في تقييم الدرجة النموذجية.

المتوسط / الوسط الحسابي The Mean

إن القياس الأكثر استخداماً في النزعة المركزية هو المتوسط الحسابي الذي يعرف بالمتوسط أو المتوسط الحسابي. إنه مجموع القيم في توزيع معين مقسم على عدد الحالات وبشكل معادلة فإنه:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{N} \qquad (5.2)$$

والتي تكتب كما يلي:

$$\bar{X} = \frac{\Sigma X}{N} \qquad (5.3)$$

حيث:

$\bar{X}$ = المتوسط الحسابي
Σ = المجموع
X = الدرجة الخام
N = عدد الحالات

وبتطبيق المعادلة (5.3) على درجات معامل الذكاء التالية، نجد إن المتوسط الحسابي هو:

112 121 115 101 119 109 100

$$\overline{X} = \frac{112+121+115+101+119+109+100}{7} = \frac{777}{7} = 111$$

لاحظ أنه في مثل هذا الحساب، لم ترتب الدرجات بأي نسق معين، فالترتيب غير ضروري لحساب المتوسط الحسابي.

ونظراً لأن الوسط هو متوسط حسابي فإنه يعتبر إحصاءة فترة. واستخدامه مناسب لبيانات الفترة أو للبيانات النسبية ولكن ليس للبيانات الاسمية أو الترتيبية.

حساب الوسط من توزيع تكراري

Computing the Mean From a Frequency Distribution

إذا تم ترتيب البيانات في توزيع تكراري فإن مجموع الدرجات يمكن حسابها بضرب كل درجة في تكرارها جمع هذه النتائج ثم تقسيم الحاصل على عدد الحالات. أما معادلة حساب المتوسط الحسابي من التوزيع التكراري فهي:

الجدول (5.3): حساب المتوسط الحسابي من توزيع تكراري لدرجات 105 طلاب في اختبار الإحصاء

X	*f*	*fX*
35	1	35
34	3	102
33	15	495
32	8	256
31	14	434
30	14	420
29	16	464
28	10	280
27	8	216
26	4	104
25	1	25
24	3	72
23	1	23
22	3	66
21	4	84
	N=105	$\Sigma fx = 3076$

$$\bar{X} = \frac{3076}{105} = 29.295$$

$$\bar{X} \quad \frac{\Sigma fx}{N} \qquad (5.4)$$

ويبين الجدول (5.4) حساب الوسط للبيانات التي عرضت في الجدول (5.1)، ومن الواضح أن استخدام المعادلة (5.4) لذلك التوزيع التكراري قد سهل حساب المتوسط الحسابي لذلك التوزيع.

مقارنة المؤشرات الثلاثة للنـــزعة المركزية

Comparison of The Three Indexes of Central Tendency

نظراً لأن المتوسط الحسابي / الوسط، إحصاءة فترية أو نسبية، فإنه مقياس أدق عموماً من الوسيط (إحصاءة ترتيبية) أو المنوال (إحصاءة اسمية). فهو يأخذ بنظر الاعتبار قيمة كل درجة. وهو الأكثر استقراراً بين المقاييس الثلاثة للنزعة المركزية من حيث أنه إذا كان عدد من العينات تحددت بشكل عشوائي من المجتمع الأصلي، فإن المتوسطات الحسابية لهذه العينات سوف

تختلف بشكل أقل فيما بينها من اختلاف الوسيطات والمنوالات. ولهذه الأسباب، كان استخدام المتوسط الحسابي أكثر شيوعاً في البحوث من المقياسين الآخرين.

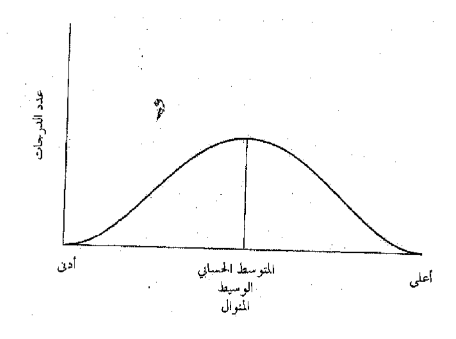
فالمتوسط الحسابي ، هو أفضل مؤشر للعمل الموحد للمجموعة برمتها. ومع ذلك، فإن الوسيط أفضل مؤشر للأداء النموذجي / النمطي. خذ مثلاً مجلس مدرسين كان الدخل السنوي لأفراده كما يأتي: 140,000 ، 60,000 ، 50,000 ، 40,000 ، 40,000 دولار. المتوسط الحسابي، 66,000 دولار هو مجموع دخولهم مقسمة على عدد الأعضاء، إلا أنه أعلى من جميع الدخول باستثناء واحد. أما الوسيط 50,000 دولار فيعطي صورة أفضل للدخل النموذجي في المجموعة.

أشكال التوزيع Shapes of Distribution

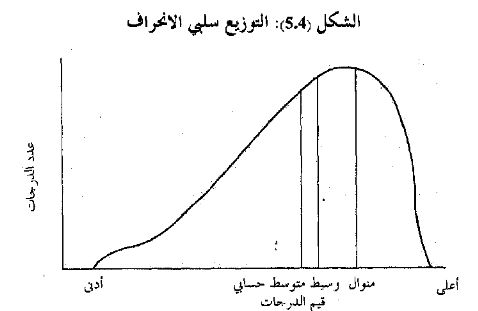
يمكن أن يكون للتوزيعات التكرارية أشكال مختلفة. فالتوزيع متناظر / متماثل حين يكون النصفان صورتين معكوستين لبعضيهما. ففي التوزيع المتماثل تتطابق قيم المتوسط الحسابي والوسيط. وإذا كان للتوزيع منوالاً واحدا بدلاً من اثنين أو أكثر فإن المؤشرات الثلاثة للنزعة المركزية سوف تتطابق كما مبين في الشكل (5.3).



الشكل (5.3): التوزيع المتناظر/ المتماثل



الشكل (5.4): التوزيع سلبي الانحراف



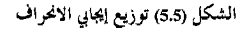
وإذا لم يكن التوزيع متناظراً فهو يوصف بأنه منحرف/ ملتو. وفي التوزيعات المنحرفة تختلف قيم مقاييس النزعة المركزية. وفي مثل هذه التوزيعات تنسحب قيمة المتوسط الحسابي، بسبب تأثرها بحجم النقاط الحدية نحو نهاية التوزيع الذي تقع فيه الدرجات المتطرفة كما هو مبين في الشكلين (5.4) و (5.5). إن تأثير القيم المتطرفة أقل على الوسيط لأن هذا المؤشر لا يتأثر بحجم الدرجات بل بمواقعها. وليس للقيم المتطرفة تأثير على المنوال لأنه لا علاقة لهذا المؤشر بأي من طرفي التوزيع. وحين يكون التوزيع منحرفاً باتجاه الطرف الأدنى، أو منحرفا بشكل سلبي، فإن المتوسط الحسابي يكون دائماً أصغر من الوسيط ويكون الوسيط، في العادة، أصغر من المنوال (الشكل 5.4).

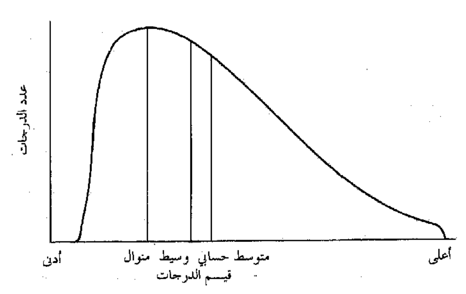
وعندما يكون التوزيع منحرفاً نحو الطرف الأعلى، أو إيجابي الانحراف، فإن المتوسط الحسابي يكون دائماً أكبر من الوسيط ويكون الوسيط في العادة أكبر من المنوال (الشكل 5.5). ويمكن تحديد انحراف/ التواء التوزيع بمقارنة المتوسط الحسابي والوسيط دون حاجة لمدرج تكراري أو مضلع.

مقاييس التغيير/ الانتشار
MEASURES OF VARIABILITY

رغم إن مؤشرات النزعة المركزية تتيح لنا وصف البيانات بدلالة قيمة المتوسط أو المقياس النموذجي/ النمطي، فإنها لا تعطينا الصورة الكاملة للتوزيع.

وقد تكون قيمتا المتوسطين الحسابين لتوزيعين متطابقتين، بينما قد تكون درجة الانتشار/ التغير لدرجاتهما ذاتها مختلفة. ففي أحد التوزيعات قد تتجمع الدرجات حول القيمة المركزية، بينما تتشتت في توزيع آخر. ولأجل الإيضاح لندرس التوزيعات التالية للدرجات:


الشكل (5.5) توزيع إيجابي الانحراف



(أ) 24، 24، 25، 25، 25، 26، 26 ويكون $\bar{X} = 175/7=25$

(ب) 16، 19، 22، 25، 28، 30، 35 ويكون $\bar{X} = 175/7=25$

إن قيمة الوسط في كلا التوزيعين هو 25، إلا إن درجة التشتت للدرجات تختلف بشكل كبير. ومن الواضح فإن الدرجات في التوزيع (أ) أكثر تجانساً من تلك التي في التوزيع (ب). ووضوحاً ، ثم حاجة لمؤشرات يمكن لها أن تصف التوزيعات بدلالة تغيرات/، تشتت الدرجات.

وتتوفر في الإحصاء عدة مؤشرات لهذا الغرض. والأربعة الأكثر شيوعاً هي المدى، الانحراف الإرباعي / الربيعي، التباين، والانحراف المعياري.

المدى Range

إن أبسط مؤشرات التغير/ التشتت هو المدى. فهو الفرق بين أعلى وأدنى الدرجات في التوزيع، ويمكن إيجاده بطرح أصغر قيمة من أعلى قيمة وإضافة (1) ويكون الأمر بشكل قانون/ معادلة كما يلي:

$$R = (X_h - X_l) + 1 \qquad (5.5)$$

حيث:

المدى = R

القيمة العليا للتوزيع = X_h

القيمة الدنيا للتوزيع = X_l

مثلاً، 15، 7 على التعاقب يمثلان المدى على التعاقب لكل من التوزيعين الآتيين:

16 14 13 12 11 10 2

15 14 13 12 11 10 9

تذكر من نقاش سابق إن كل درجة في التوزيع تمثل فترة من منتصف المسافة بين الدرجة وأقل درجة تسبقها حتى منتصف المسافة بين تلك النقطة وأعلى درجة تتلوها. مثلاً العدد 4 يمثل النقطة الوسط للفترة 3.5 حتى 4.5. وهكذا لإيجاد المدى نأخذ هذا الحد الحقيقي الأعلى، لأعلى درجة ناقصاً الحد الحقيقي الأدنى، لأدنى درجة.

في المثال الأول أعلاه، يتم إيجاد المدى بطرح (1.5) (أدنى حد لأقل درجة) من (16.5) (أعلى حد لأعلى درجة) والمساوي لـ (15)، في المثال الثاني يكون المدى (15.5 - 8,5) أو (7). ومن السهل استخدام المعادلة (5,5) وطرح العدد الأدنى من العدد الأعلى وإضافة (1)، في المثال الأول (15=1+2-16) وفي الثاني (7=1+9-15).

المدى مؤشر غير موثوق للتشتت لأنه يستند فقط على قيمتين، العليا والدنيا. وكما نرى من التوزيع أعلاه، فإنه مؤشر غير مستقر لطبيعة انتشار القياسات حول القيمة المركزية. ولهذا السبب فإن استخدام المدى محدد أساساً بأهداف استقصائية. ففي بعض التقارير البحثية، تتم الإشارة إلى مدى التوزيعات إلا إن هذه المرجعيات تستخدم عادة بالارتباط مع قياسات أخرى للتشتت مثل الانحراف الربيعي والانحراف المعياري.

الانحراف الربيعي Quartile Deviation

إن الانحراف الربيعي (QD) هو نصف الفرق بين الربيعي الأدنى والربيعي الأعلى في التوزيع. والربيع الأعلى (Q_3) هو النقطة في التوزيع التي تقع دونها 75% من الحالات. أما الربيع الأدنى (Q_1) فهو النقطة التي تقع دونها 25% من الحالات. إن الربيع الأعلى يشار إليه كذلك بأنه المئيني الخامس والسبعين بينما يشار إلى الربع الأدنى بالمئيني الخامس والعشرين.

إن إجراء إيجاد Q_1، Q_3 مماثل لذلك المستخدم لإيجاد الوسيط، والوسيط هو في الواقع الربيع الثاني، وفي حالة Q_3 تصبح المعادلة/ القانون:

$$Q_3 = L + \left(\frac{\frac{3N}{4} - cfd}{fw} \right) \qquad (5.6)$$

وتكون المعادلة بالنسبة لـ Q_1

$$Q_1 = L + \left(\frac{\frac{N}{4} - cfd}{fw} \right) \qquad (5.7)$$

حيث:

Q_3 = الربيع الأعلى

Q_1 = الربيع الأدنى

L = الحد الأدنى للفاصلة التي يقع فيها الربيع

N = عدد الحالات في التوزيع

Cfd = التكرار المتراكم تحت الفترة التي تحتوي الربيع

Fw = تكرار الحالات في الفترة الحاوية على الربيع

i = حجم/ طول الفترة

وعند إيجاد الربيعين الأول والثالث فإن الانحراف الربيعي يمكن حسابه كما يلي:

$$QD = \frac{Q_3 - Q1}{2}$$

فمثلاً إذا كان الربيع الأعلى والربيع الأدنى في توزيع الدرجات هو (35) و (15) على التعاقب فإن قيمة الانحراف الربيعي سيكون (10) كما مبين هنا:

$$QD = \frac{35 - 15}{2} = 10$$

الانحراف الربيعي، مقياس للانتشار خلال النصف الوسط للتوزيع، وعليه فأنه لا يتأثر بالدرجات العالية جداً أو المنخفضة جداً.

ورغم إن الانحراف الربيعي، مثل المدى، يستند على قيمتين فقط في التوزيع (Q_3, Q_1) فإنه قد يعد مقياساً أكثر فائدة للتشتت من المدى. ويعزى ذلك إلى إن (Q_3, Q_1) هي أكثر استقراراً من القيم العليا والدنيا التي يستند إليها المدى.

وينتمي الانحراف الربيعي إلى العائلة الإحصائية نفسها كالوسيط لأنه إحصاءة ترتيبية، وهو غالباً ما يستخدم مع الوسيط، كما أنه يدعى كذلك (شبه المدى الربيعي). وكما الوسيط، فإنه مفيد بشكل خاص حين تريد مقياساً لا يتأثر بنقاط متطرفة قليلة. فإذا كان الانحراف الربيعي للمداخيل العائلية في القرية (أ) هو (10,000) دولار وفي القرية (ب) (8000) دولار، فإننا سنعرف أن المداخيل العائلية في (أ) هي أكثر اختلافاً حسب قياسها بالانتشار عبر منتصف التوزيع. ولو فازت عائلة واحدة في القرية (ب) بجائزة يا نصيب كبيرة فإنها لن تغير، أو تغير القليل في الوسيط أو الانحراف الربيعي، ومع ذلك، فإن هذا سيزيد من المتوسط الحسابي والانحراف المعياري بشكل كبير.

التباين والانحراف المعياري Variance and Standard Deviation

التباين والانحراف المعياري، هما أكثر المقاييس فائدة للتشتت/ الانتشار. وكلاهما يستند على درجات الانحراف - وهي درجات تبين الفرق بين درجة خام والوسط. أما معادلة درجة الانحراف فهي ($x = X - \bar{X}$). فالدرجات الخام دون المتوسط الحسابي سيكون لها درجات انحراف سلبية، أما الدرجات الخام فوق المتوسط الحسابي فسيكون لها درجات انحراف إيجابية. وبالتعريف فأن مجموع درجات الانحراف في توزيع معين هي (صفر) دائماً. وهكذا لو أراد المرء أن يستخدم درجات الانحراف في حساب قياسات التشتت فلابد أن يجد طريقة للإحاطة بحقيقة أن $\Sigma x = 0$. والأسلوب المستخدم هو تربيع كل درجة انحراف حيث ستصبح أعداداً موجبة. وإذا جمعنا مربعات الانحرافات وقسمناها على عدد الدرجات فسوف نحصل على متوسط حسابي لمربعات الانحرافات عن المتوسط الحسابي أو (التباين)، وبشكل حسابي فإن (التباين) هو

$$\sigma^2 = \frac{\sum x^2}{N} \qquad (5.9)$$

حيث:

σ^2 = التباين

Σ = المجموع

$\bar{x}$ = الوسط الحسابي

x = انحراف كل درجة x عن المتوسط الحسابي ($X - \bar{X}$) والمعروفة بدرجة الانحراف

N = عدد الحالات في التوزيع

ونظراً لأنه جرى تربيع كل من درجات الانحراف فإنه من الضروري التعبير عن التباين بوحدات هي مربعات الوحدات الأصلية للقياس. مثلاً، قد نجد أن التباين في أطوال الأطفال في صف معين هو (9) بوصات مربعة. وسيدلنا هذا على أن الصف أكثر اختلافاً في الطول من الصف ذي التباين (4) بوصات مربعة وأكثر تجانساً من الصف ذي (16) بوصة مربعة.

ويفضل التربويون، في أغلب الحالات، مؤشراً يلخص البيانات في وحدة القياس ذاتها كالبيانات الأصلية. إن الانحراف المعياري (σ) وهو الجذر التربيعي للتباين يقدم مثل هذا المؤشر. وبصورة واسعة فإنه المقياس الأكثر استخداماً للتشتت. وبالتعريف فإن الانحراف المعياري هو الجذر التربيعي للوسط الحسابي لمربعات درجات الانحراف. وبإعادة كتابة التعريف باستخدام الرموز نحصل على:

$$\sigma = \sqrt{\frac{\sum x^2}{N}} \qquad (5.10)$$

ولغرض الإيضاح أدرس الجدول (5.4). فالعمود (1) في هذا الجدول يبين توزيع درجات اختبار عشرة أشخاص. والمتوسط الحسابي لهذا التوزيع هو (6). ويعرض العمود (2) الانحرافات لكل درجة من الدرجات، مثلاً انحراف النقطة (8) عن والوسط هي (+2) وانحراف الدرجة (5) عن المتوسط الحسابي هي (−1) وهكذا. ويبين العمود (3) مربعات كل من درجات الانحراف هذه. ومجموع مربعات درجات الانحراف هو (90). وبوضع هذه القيمة في المعادلة (5.10) وتقسيمها على (10) فإن عدد الحالات التي نصل إليها هو (9) الذي هو المتوسط الحسابي لمربعات درجات الانحراف. إن الجذر التربيعي لهذه القيمة هو (3) الذي هو الانحراف المعياري لهذا التوزيع.

الجدول (5.4): حساب الانحراف المعياري

(1) X	(2) $x = X - \overline{X}$	(3) $x^2 = (X - \overline{X})^2$
10	+4	+16
9	+3	+9
9	+3	+9
8	+2	+4
7	+1	+1
6	0	0
5	-1	+1
3	-3	+9
2	-4	+16
1	-5	+25
$\Sigma X = 60$		$\Sigma x^2 = 90$

$$\overline{X} = \frac{60}{10} = 6, \qquad \sigma = \sqrt{\frac{90}{10}} = \sqrt{9} = 3$$

يعد الإجراء السابق مناسباً حين يكون وسط التوزيع هو عدد صحيح أو مقرب لعدد صحيح، ولكن ذلك لا يحدث في أغلب الحالات. وعليه فإن المعادلة التالية قد طورت لغرض استبعاد العمل الممل مع درجات الانحراف الكسرية. ويعطي استخدام هذه الصيغة النتيجة ذاتها بجهد أقل. وهكذا فإنه يوصي بأن يستخدم الطلبة دائماً هذه الصيغة/ المعادلة لحساب الانحراف المعياري:

$$\sigma = \sqrt{\frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N}} \qquad (5.11)$$

حيث:

σ = الانحراف المعياري

ΣX^2 = مجموع مربعات الدرجات (أي يحسب أولاً مربع كل درجة وتجمع المربعات)

$(\Sigma X)^2$ = مربع مجموع الدرجات (أي تجمع الدرجات أولا ويحسب مربع المجموع)

N = عدد الحالات

يبين الجدول (5.5) حساب الانحراف المعياري للبيانات في الجدول (5.4) باستخدام الصيغة (5.11). العمود الأول في هذا الجدول يبين الدرجات ومجموعها. ويبين العمود الثاني مربع كل درجة ومجموع هذه المربعات.

الجدول (5.5) حساب الانحراف المعياري باستخدام الصيغة/ القانون (5.11)

X	X^2	
10	100	$\sigma=\sqrt{\dfrac{450-\dfrac{(60)^2}{10}}{10}}$
9	81	
9	81	
8	64	$=\sqrt{\dfrac{450-360}{10}}$
7	49	
6	36	
5	25	$=\sqrt{9}$
3	9	$=3$
2	4	
1	1	
$\Sigma X=60$	$\Sigma X^2=450$	

وتبين بقية الجدول تطبيق الصيغة (5.11) لإيجاد قيمة الانحراف المعياري. لاحظ أن القيمة الناتجة هي نفسها كالتي وجدت من تطبيق الصيغة (5.10) على نفس البيانات.

وينتمي الانحراف المعياري إلى نفس العائلة الإحصائية كالمتوسط الحسابي، فهو مماثل للمتوسط الحسابي بكونه إحصاءة فترية أو نسبية، ويستند حسابه على حجم الدرجات الإفرادية في التوزيع. أنه بشكل عام القياس الأكثر استخداماً للتغير/ التشتت ويترافق في الاستخدام مع المتوسط الحسابي.

الدرجات المعيارية STANDARD SCORES

نرغب في الغالب بإجراء مقارنات بين المواقع النسبية لفرد واحد في اختبارين مختلفين. ومن الممكن إجراء ذلك بشكل معقول إذا كان للاختبارين نفس المتوسطين الحسابيين ونفس الانحرافين المعياريين، إلا إن ذلك نادراً ما يحدث عملياً. وللتغلب على هذه الصعوبة يمكن أن نترجم القياسات إلى درجات معيارية وهي درجات محمولة إلى توزيعات لها متوسطات انحرافية وانحرافات معيارية ذات قيمة معيارية. والدرجة المعيارية الأوسع استخداماً التي تلعب دوراً مهماً في التحليلات الإحصائية هي الدرجة (Z) التي تعرف بمسافة الدرجة عن المتوسط الحسابي مقاسة بوحدات الانحراف المعياري. والمعادلة لاستخراج الدرجة (Z) هي:

$$Z = \frac{X - \overline{X}}{\sigma} = \frac{x}{\sigma} \qquad (5.12)$$

حيث:

X = الدرجة الخام
$\overline{X}$ = المتوسط الحسابي للتوزيع
σ = الانحراف المعياري للتوزيع
x = انحراف الدرجة ($X - \overline{X}$)

وبتطبيق المعادلة هذه فإن الدرجة التي فوق المتوسط الحسابي بانحراف معياري واحد، تصبح (Z= +1) والدرجة بانحراف معياري واحد دون المتوسط الحسابي تصبح (Z= -1) وهكذا. أما النقطة المساوية للوسط فستكون لها قيمة الدرجة (Z) التي تساوي الصفر. ولغرض الإيضاح افترض إن علامة تلميذ في اختبار علم النفس هي (72) حيث المتوسط الحسابي للتوزيع هو (78) والانحراف المعياري يساوي (12). وأفترض كذلك إن ذات التلميذ قد حصل على درجة (48) في اختبار الإحصاء حيث المتوسط الحسابي هو (51) والانحراف المعياري (6). فإذا استبدلنا هذه الأعداد برموز مناسبة في المعادلة (5.12) لأمكننا أن نستنتج النقطة (Z) لكل اختبار:

علم الإحصاء

$$Z_2 = \frac{48 - 51}{6} = -0.50$$

علم النفس

$$Z_1 = \frac{72 - 78}{12} = -0.50$$

إن كلا الدرجتين المعياريتين تنتميان إلى التوزيع (Z) حيث بالتعريف يكون المتوسط

الحسابي صفراً والانحراف المعياري (1)، وعليه فإنها موضع مقارنة بشكل مباشر. ومن الواضح في هذا المثال فإن الدرجة (72) في اختبار علم النفس والدرجة (48) في اختبار علم الإحصاء متكافئتان – أي إن كلا الدرجتين تشيران إلى ذات المستوى النسبي للتحصيل. وبعبارة أخرى، إن موقف الطالب الذي حصل على هذه الدرجات هو ذاته في كلا الاختبارين لدى مقارنته مع أداء الطلبة الآخرين. وسيكون من الصعب إجراء مثل هذه المقارنة دون استخدام أسلوب الدرجة (Z). دعنا نستخدم مثالاً أخر: أفترض أن تلميذاً أخذ نفس الاختبارات وحصل على درجة (81) في اختبار علم النفس ودرجة (53) في اختبار علم الإحصاء وكما في السابق، فإنه من الصعب مقارنة هذه الدرجات الخام كي نتبين في أي اختبار كان التلميذ قد أدى الاختبار بشكل أفضل . وتحويل الدرجات إلى درجات (Z) يجعل المقارنة سهلة. وباستخدام المعادلة (5.12) نجد إن قيم (Z_1) و (Z_2) في هذه الحالة هي كما يلي:

علم الإحصاء

$$Z_2 = \frac{54-51}{6} = +0.33$$

علم النفس

$$Z_1 = \frac{81-78}{12} = +0.25$$

تبين هذه النتيجة إن العلامة (53) في اختبار الإحصاء توضح فعلاً أداءاً أفضل نسبياً من العلامة (81) في اختبار علم النفس. وبالمقارنة مع الطلبة الآخرين فأن هذا الطالب قد أدى أفضل بالإحصاء منه في علم النفس.

إن إحدى عيوب الدرجات (Z) هي إن علينا أن نتعامل مع قيم سلبية وكسور عشرية. وللتغلب على هذه الصعوبات يمكن أن نحول درجات (Z) إلى مقياس درجات معيارية أخرى لا ينطوي على أرقام سلبية أو عشرية. إن أحد الأمور المألوفة هو توزيع الدرجة (T) الذي له متوسط حسابي (50) وانحراف معياري (10). ولتحويل درجات (Z) إلى درجات (T) نضرب قيمة (Z) في (10) ونضيف (50). أما معادلة الدرجة T فهي:

$$T = 10z + 50 = 10\left(\frac{X-\overline{X}}{\sigma}\right) + 50 \qquad (5.13)$$

افترض أن درجة تلميذ في اختبار الأسبانية هي (21). وإذا اعتبرنا أن الوسط الحسابي لدرجات هذا الاختبار (27) وإن الانحراف المعياري (6) فالدرجة (Z) ستكون 6/(21-27)، والتي يمكن إدراجها مباشرة في معادلة الدرجة (T) كما يلي:

$$T = 10\left(\frac{21-27}{6}\right) + 50 = 40$$

إن تحويل نقاط (Z) إلى درجات (T) لا يمكن للمرء من التعامل مع جميع الأعداد، فحسب، بل إنها تتجنب المضامين النفسية غير الملائمة لوصف أداء الأشخاص بأعداد سالبة.

أما المدرسون الذين يرغبون في مقارنة مواقف تلاميذهم في اختبارات متعاقبة أو لإضافة جميع الدرجات التي حصلوا عليها في اختبارات مختلفة في المقرر الدراسي ذاته للحصول على توزيع عام، فبوسعهم يمكن أن يحولوا الدرجات الخام للتلاميذ إلى درجات (z) أو (T) بغية إعطاء وزن مساوٍ لكل مجموعة من الدرجات. فجمع الدرجات أو حساب متوسطاتها التي تعود إلى توزيعات مختلفة ولها متوسطات حسابية مختلفة وانحرافات معيارية مختلفة دون تحويلها إلى نوع من النقطة المعيارية هو أمر غير مبرر إحصائياً.

إضافة إلى (T) هناك تحويلات أخرى لتوزيعات الدرجة المعيارية. ولتحويل توزيع نقاط على توزيع معياري جديد فإنه يلزم فقط، ضرب درجة (Z) بالانحراف المعياري المرغوب ثم إضافة الوسط الحسابي المرغوب. والصيغة العامة هي كما يلي:

$$A=\mu_A+\sigma_A(Z) \quad (5.13a)$$

حيث:

A = الدرجة المعيارية على المقياس الجديد
μ_A = المتوسط الحسابي للمقياس المعياري الجديد
σ_A = الانحراف المعياري للمقياس المعياري الجديد

فمثلاً، كل جزء من سجل اختبار التخرج (GRE) له متوسط حسابي من (500) وانحراف معياري بمعدل (100) لأجل توزيعه المحول. فإذا كان موقفك هو انحراف معياري (1.5) فوق المتوسط الحسابي، أي (Z = 1.5) في الجزء الشفوي من (GRE) فإن درجتك ستكون 650، أي [500 + 100 (1.5)]. وإذا كان مقدار درجتك هي (500) فإن نتيجتك تماماً على المتوسط الحسابي.

إن مقياس ستانفورد – بينيه للذكاء هو مقياس درجات معيارية ذات متوسط حسابي هو (100)، وانحراف معياري هو (16)، ومن ثم لتحويل درجة خام على مقياس ستانفورد – بينيه للمقياس المعياري، يجب على المرء أولاً أن يحدد قيمة (Z) ثم يضربها في (16) ويضيف (100).

إن تحويل مجموع درجات إلى درجات معيارية لا يغير شكل التوزيع الأصلي. فإذا كان توزيع الدرجات منحرفاً فإن الدرجات المعيارية المشتقة ستنتج توزيعاً منحرفاً. وفقط، إذا كان التوزيع الأصلي اعتيادياً فأن النقاط المعيارية سوف تتمخض عن توزيع اعتيادي / نظامي.

المنحنى العادي/ الاعتدالي/ المعياري THE NORMAL CURVE

لقد وجد أن توزيع العديد من القياسات البدنية والنفسية يأخذ شكل الجرس عند تخطيطه كمضلع تكراري. مثلاً، إذا قمنا بقياس الأولاد الأمريكان في عيد ميلادهم العاشر، سنجد أن

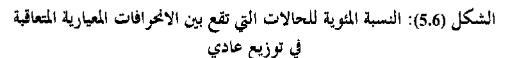
العديد من الأولاد يبلغ طولهم قريبا من المتوسط الحسابي، وقلة من الأولاد ممن هم أعلى قليلاً أو أدنى قليلاً من المتوسط الحسابي . وكلما ابتعدنا من المتوسط الحسابي، سنجد أولاداً أقل في كل طول. إن المضلع الذي يبين هذا التوزيع يشبه تماماً مضلعاً نظرياً يعرف باسم (المنحنى العادي). إن جداول الإحداثيات الرأسية ومناطق المنحنى العادي، قد تم حسابها. وتتضح المناطق في الجدول (A-1) في الملحق. ويوضح هذا المضلع الفرضي التكرارات المتوقعة (أي النظرية) لجميع درجات (Z) المحتملة. وهو يوضح أن درجات (Z) قرب الصفر سيتوقع لها أن تحدث بصورة أكثر من قيم الدرجة (Z) الأخرى، وكلما كنا أبعد من نقطة (Z) الصفرية، قلّ توقع تكرار حدوثها.

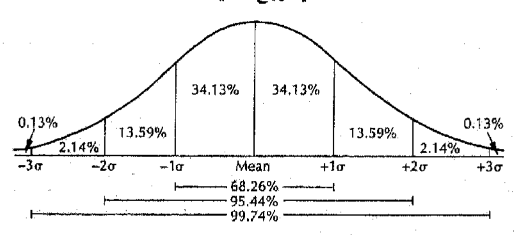
ونظراً لأن العديد من التوزيعات التي تحدث بشكل طبيعي تشبه المنحى العادي، فإن هذا النموذج النظري قد برهن على فائدة عالية، وكلما كانت البيانات الفعلية معروفة أو اعتقدنا أنها تشابه المنحي العادي في التوزيع، كان بوسعنا أن نستنتج العديد من التقديرات المفيدة، من الخواص النظرية للمنحنى الاعتيادي/ العادي (أو النظامي أيضاً).

يعتبر المنحنى العادي توزيعاً متناظراً للقياسات مع ذات العدد من الحالات على مسافات محددة دون المتوسط الحسابي وفوقه. أما متوسطها الحسابي فهو النقطة التي تقع دونها 50 بالمائة من الحالات كما تقع 50 بالمائة الحالات فوقها. أما الوسيط والمنوال لمثل هذا التوزيع فقيمتاهما متطابقتان مع الوسط الحسابي. وفي المنحنى العادي تتركز أغلب الحالات قرب المتوسط الحسابي. أما تكرار الحالات فيتناقص كلما ابتعدنا من الوسط في أي من الاتجاهين. وقرابة 34 في المئة من الحالات في التوزيع العادي تقع بين المتوسط الحسابي وبين انحراف معياري واحد فوق دون المتوسط الحسابي. أما المنطقة بين انحراف معياري واحد وانحرافين معياريين من المتوسط الحسابي على كل جانب من التوزيع، فيحتوي على حوالي 14 في المئة من الحالات، وفقط حوالي 2 في المئة من الحالات تقع بين انحرافين أو ثلاثة معيارية من المتوسط الحسابي. وفقط حوالي 0.1 في المئة من الحالات تقع فوق أو تحت الانحرافات المعيارية الثلاثة من المتوسط الحسابي. ويتضح ذلك في الشكل (5.6). ومن الممكن تحديد النسبة المئوية للحالات دون أو فوق كل درجة (Z) في التوزيع العادي، وذلك بالرجوع إلى الجدول (A-1) في الملحق الذي يعطى مناطق المنحنى العادي. العمود (1) من الجدول يحتوي على قيم (Z) المختلفة. أما العمود (2) فيعطي المنطقة تحت المنحنى بين المتوسط الحسابي وكل قيمة لـ (Z). ويوضح العمود (3) المنطقة الباقية من كل قيمة لـ (Z) حتى نهاية المنحنى.

وعليه، فإن مجموع المنطقتين في العمود (2) والعمود (3) يبلغ (0.5000). خذ كمثال قيمة (Z=+0.70). المنطقة بين قيمة (Z) هذه والمتوسط الحسابي يمكن إيجادها في العمود (2)، وهي (0.2580). ويوضح هذا العدد أن حوالي 26 في المئة من الحالات تقع بين هذه القيمة (Z) هذه، والوسط الحسابي للتوزيع. ولأن المتوسط الحسابي للتوزيع العادي يتطابق مع الوسيط

فإن 50 في المئة من الحالات تقع دون الوسيط. فإذا أضفنا (0.50) إلى (0.2580) فإن النتيجة تبين أن بوسعنا أن نتوقع أن (75.8 في المئة) من الحالات تقع دون (Z = 0.70). ويشير العمود (3) إلى أن 24.2 في المئة من الحالات تقع فوق Z = 0.70.


الشكل (5.6): النسبة المئوية للحالات التي تقع بين الانحرافات المعيارية المتعاقبة في توزيع عادي



وينعكس هذا الأجراء عندما تكون قيمة (Z) سالبة. أفترض أننا نريد إيجاد النسبة المئوية للحالات دون القيمة 0.70 - = Z. إن المنطقة بين المتوسط الحسابي والدرجة 0.70 - = Z. إن هي (0.2580) بدلالة النسبة المئوية فإنها تشير إلى (25.8 في المئة) من الحالات. وبطرح (25.8) من (50) نحصل على (24.2). وتوضح هذه النتيجة أن حوالي 24 في المئة من الدرجات فقط، تقع دون قيمة 0.70- = Z في توزيع عادي. ويمكن إيجاد هذه القيمة أيضاً في العمود (3) من الجدول الذي يعطي فيه (0.2420) للدرجة Z = 0.2420. أما نسبة الدرجات فوق 0.70- = Z فهي (100) ناقص (24.2) أو 75.8 في المئة.

إن المناطق تحت المنحى العادي بين درجتين لـ (Z) تنطبق بالضبط، فقط، على التوزيع العادي. أما تطبيق خواص المنحى العادي على التوزيعات المنحرفة، فينتج عنها، في أحسن الأحوال، قيماً تقريبية فحسب. وهكذا فإن من الضروري أن يختبر المرء شكل التوزيع قبل تطبيق هذه الخواص.

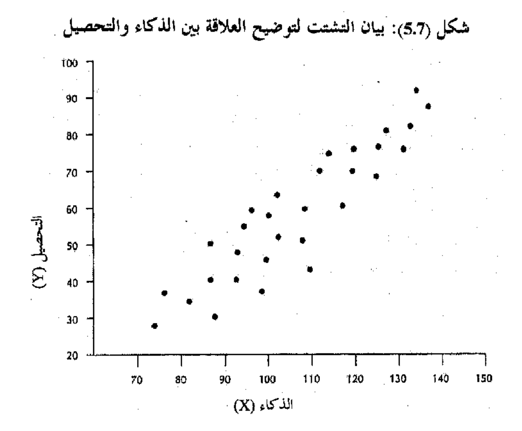
الارتباط CORRELATION

إن نقاشنا للأساليب الإحصائية كان يخص وصف التوزيعات المفردة للدرجات، والآن سنناقش طريقة إيضاح العلاقة بين أزواج من الدرجات.

بيان التشتت/ الانتشار(*) Scattergrams

تعرف الأساليب الإحصائية لتحديد العلاقات بين أزواج الدرجات (بالإجراءات الارتباطية). ونموذجيا، فإن القياسات الخاصة بمتغيرين تتوفر لكل فرد من المجموعة، ويحدد المرء ما إذا كانت هناك علاقة بين هذه القياسات الزوجية. وتبين الإجراءات الارتباطية، المدى الذي يرتبط فيه التغير في متغير واحد مع التغير في متغير آخر. فمثلاً، نحن نعلم أن التحصيل والذكاء مترابطان، ولذا نتوقع من التلاميذ ذوي مستوى الذكاء العالي أن يحصلوا على درجات أعلى من المعدل في اختبارات التحصيل. والطريقة البسيطة لإيضاح هذه العلاقة هي بتعيين درجات اختبار الذكاء ودرجات اختبار التحصيل لعدد من الأفراد في جدول ذي بعدين يدعى بيان التشتت. ويتم تعيين الدرجات لمتغير واحد على محور أفقي مع أدنى عدد على اليسار وأعلى عدد على اليمين. أما درجات المتغير الآخر فيتم تعيينها على محور عمودي، مع الأدنى إلى الأسفل والأعلى إلى القمة. ويتضح موقع كل فرد في الاختبارين بعدئذ بنقطة واحدة في بيان التشتت. فدرجات اختبار التحصيل لـ 30 تلميذا من الصف العاشر تم تعيينها إزاء درجات اختبار الذكاء لديهم في الشكل (5.7). ويكشف تفحص هذا الشكل إن هناك ميلا لأن تكون درجات التحصيل عالية حين تكون درجات اختبارات الذكاء عالية.

شكل (5.7): بيان التشتت لتوضيح العلاقة بين الذكاء والتحصيل



(*) بيان التشتت/ الانتشار Scattergrams: يشير "البيان" رياضيا إلى مجموعة النقاط البيانية التي تحددها علاقة ما، وهناك من يسمية "مبيان التشتت" – (المراجع)

ويتيح لنا بيان التشتت رؤية كل من اتجاه وقوة العلاقة. فالاتجاه يشير إلى ما إذا كانت العلاقة إيجابية أم سالبة. ففي الشكل (5.7) تشكل النقاط نمطاً يبدأ من أسفل اليسار إلى أعلى اليمين حيث الدرجات الدنيا لمتغير (الذكاء) ترتبط مع الدرجات الدنيا للمتغير الآخر (التحصيل)، والدرجات العليا لمتغير تقترن بالدرجات العليا للمتغير الآخر. (واتفاقاً، يتم تعيين درجات المتغير المستقل (X) على طول محور أفقي ودرجات المتغير التابع (Y) تُعين على محور عمودي) ويقال إن مثل هذه العلاقة بين المتغيرين موجبة لأن الدرجات العالية تقترن بالدرجات العالية والدرجات الدنيا مع الدرجات الدنيا.

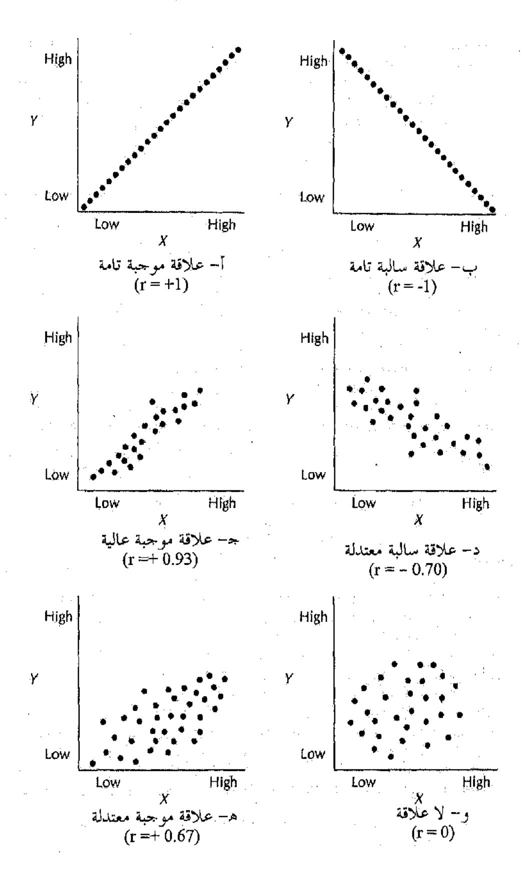
ولا تكون العلاقة بين المتغيرين موجبة دائماً. فبعض المتغيرات ترتبط بشكل سلبي. فمثلاً، وجد إن معدل الولادات والمستوى الاجتماعي الاقتصادي يرتبطان بشكل سلبي، أي إن معدل الولادة يتناقص كلما زاد المستوى الاجتماعي الاقتصادي. فمع العلاقة السلبية ترتبط الدرجات العالية لمتغير معين مع الدرجات المتدنية للمتغير آخر، وتبدأ نقاط بيان التشتت من اليسار الأعلى إلى اليمين الأدنى.

ويكشف بيان التشتت للدرجات (Z) قوة العلاقة بين المتغيرات. فإذا كانت النقاط (dots) في بيان التشتت تشكل شريطاً ضيقاً بحيث إذا رسم خط مستقيم عبر الشريط تكون هذه النقاط قرب الخط. فإن هناك علاقة قوية بين المتغيرين. وعلى أية حال إذا كانت النقاط (dots) في بيان تشتت درجات (Z) تنتشر بشكل واسع فإن العلاقة بين المتغيرين ضعيفة نسبياً. ويوضح بيان التشتت في الشكل (5.8) علاقات متعددة موجبة وسالبة وقوية وضعيفة.

معاملات الارتباط Correlation Coefficients

لقد تطورت المؤشرات الإحصائية التي توضح كلاً من الاتجاه (السالب أو الموجب) وقوة العلاقة بين المتغيرات. وتدعى هذه المؤشرات (معاملات الارتباط). إن حساب معامل الارتباط بين متغيرين ينتج قيمة تتراوح من (1-) إلى (1+). فمعامل ارتباط (1-) يبين علاقة سالبة تامة، بينما تشير القيمة (1+) إلى علاقة موجبة تامة، والنقطة الوسطى لهذا المدى وهي (صفر) تشير إلى عدم وجود أية علاقة على الإطلاق. إن الارتباط الموجب التام ينتج حين تكون الدرجة (Z) لكل شخص على متغير معين مماثلة في الحجم والإشارة للدرجة (Z) على المتغير الآخر. ومن ناحية أخرى، ينتج الارتباط السالب التام حين تكون كل درجة (Z) لشخص معين مماثلة في الحجم ومتعاكسة في الإشارة. أما الارتباط (صفر) فينتج حين لا تكون هناك مثل هذه الاتجاهات، أي حين لا ترتبط المواقع لمتغير مع المواقع للمتغير الآخر. ويوضح معامل الارتباط القريب من الوحدة، (1-) أو (1+) درجة عالية من العلاقة. وتسمح مثل هذه العلاقات العالية بتنبؤات دقيقة حول متغير معين على أساس المعلومات عن الآخر. ويعد معامل الارتباط السالب مناسباً للتنبؤ كمعامل الارتباط الموجب.

الشكل (5.8): مقاييس التشتت لقيم r المنتخبة



يوضح الشكل (5.8)، معامل ارتباط كل مجموعة من البيانات حسب الدرجة (Z). لاحظ أنه عندما يكون الارتباط تاماً فأن جميع النقاط تقع على خط مستقيم. وكلما كان معامل الارتباط قريباً من (الصفر) كان انحراف النقاط عن الخط المستقيم أكبر. في المثال (و)، الذي يوضح الارتباط (صفر) تتشتت النقاط على سطح الشكل، ولا تأخذ أي شكل في أي اتجاه.

إن معاملات الارتباط في المقاييس التربوية والنفسية، وبسبب تعقيد هذه الظواهر، قلما تصل إلى النقاط القصوى (1+) و (1−) . فبالنسبة إلى هذه المقاييس يعتبر كل معامل أكبر من (0.90) أو أصغر من (0.90−) عالياً جداً.

الارتباط التتابعي/ ارتباط حاصل ضرب العزوم

The Product Moment Correlation

إن معامل الارتباط التتابعي الذي طوره الإحصائي الإنكليزي كارل بيرسون وسمي بمعامل ارتباط بيرسون (*r* Pearson) هو مؤشر الارتباط الأكثر شيوعاً في الاستخدام. فمعامل الارتباط هذا يستخدم حين يكون المقياس فتريا أو نسبيا.. ويعرف بأنه المتوسط الحسابي لحاصل ضرب الدرجات (Z)، أي تضرب كل درجة (Z) لشخص ما على المتغير (X) بدرجته (Z) على المتغير الآخر (Y) تجمع حواصل ضرب درجات (Z) الزوجية ويقسم الناتج على عدد الأزواج. وتعريفا، فإن معامل ارتباط بيرسون (r) هو:

$$r = \frac{\sum z_x z_y}{N} \qquad (5.14)$$

حيث:

r = معامل بيرسون للارتباط التتابعي
$\sum z_x z_y$ = مجموع حواصل ضرب الدرجات Z
N = عدد أزواج الدرجات

وبسبب الطريقة التي تعرف فيها الدرجات (Z) رياضياً، فحينما يكون لكل فرد نفس الدرجة (Z) على (X) و على (Y) فإن مجموع حواصل Z_x . Z_y سيكون مساوياً لعدد الأزواج وإن المتوسط الحسابي لحاصل ضرب الدرجات (Z) (معامل ارتباط بيرسون) سيكون (1.00+). وإن كان هناك تطابق موجب تام بين درجات (Z) فأن معامل ارتباط لحظة النتائج سيكون (1+) أما إذا كانت درجات (Z) متطابقة عددياً ولكن بإشارة مضادة فإن حاصل الضرب سيكون سالباً ومعامل بيرسون سيكون (1−). وتقع في كلا الحالتين جميع النقاط على خط مستقيم عند رسم بيان التشتت. دعنا الآن نستخدم المعادلة (5.14) لحساب الارتباط بين درجات 14 شخصاً في اختبارين، (X) (الإحصاء الوصفي) و (Y) (للإحصاء الاستدلالي)،

كما مبين في الجدول (5.6).

الجدول (5.6): حساب معامل ارتباط بيرسون بين مجموعتي درجات (X, Y)

(1) الأفراد	(2) X- الدرجات	(3) x	(4) x^2	(5) Y- الدرجات	(6) y	(7) y^2	(8) z_x	(9) z_y	(10) z_xz_y
1	18	+3	9	28	4	16	+1.5	+1	+1.50
2	18	+3	9	30	6	36	+1.5	+1.5	+2.25
3	17	+2	4	30	6	36	+1	+1.5	+1.50
4	17	+2	4	26	2	4	+1	+0.5	+0.50
5	16	+1	1	28	4	16	+0.5	+1	+0.50
6	16	+1	1	24	0	0	+0.5	0	0
7	15	+0	0	22	-2	4	0	-0.5	0
8	15	+0	0	20	-4	16	0	-1	0
9	14	-1	1	26	2	4	-0.5	+0.5	-0.25
10	14	-1	1	22	-2	4	-0.5	-0.5	+0.25
11	13	-2	4	24	0	0	-1	0	0
12	13	-2	4	18	-6	36	-1	-1.5	+1.50
13	12	-3	9	20	-4	16	-1.5	-1	+1.50
14	12	-3	9	18	-6	36	-1.5	-1.5	+2.25
	210		56	336		224			11.50

$$\bar{X} = \frac{\Sigma X}{N} = \frac{210}{15} = 15 \qquad \bar{Y} = \frac{\Sigma Y}{N} = \frac{336}{14} = 24$$

$$\sigma_x = \sqrt{\frac{\Sigma x^2}{N}} = \sqrt{\frac{56}{14}} = 2 \qquad \sigma_y = \sqrt{\frac{\Sigma y^2}{N}} = \sqrt{\frac{224}{14}} = 4$$

يعرض العمودان (2) و (3) الدرجات الخام للأشخاص (X) ودرجات الانحراف (X) على التعاقب في اختبار الإحصاء الوصفي. ويعرض العمودان (5) و (6) الدرجات الخام للأشخاص (Y) ودرجات الانحراف (Y) في اختبار الإحصاء الاستدلالي.

يوضح العمودان (4) و (7) مربع انحراف الدرجات المستخدمة في حساب الانحرافات المعيارية. أما الأعمدة (8) و (9) فتبين درجات (Z) لكل من درجات (X) و (Y) المحسوبة باستخدام المعادلة (5.12). أما العمود (10) فيبين قيم حواصل $Z_x \cdot Z_y$. إن مجموع حواصل الضرب هذه يساوي (11.50). ويصبح من الممكن الآن إيجاد معامل ارتباط بيرسون بين مجموعتي الدرجات بتطبيق المعادلة (5.14):

$$r=\frac{11.50}{14}=+.85$$

إن عملية تحويل الدرجات إلى درجات (Z) تصبح مملة عند استخدام عدد كبير من الحالات. ومن الممكن استبعاد هذه الخطوة والعمل مباشرة مع الدرجات الخام باستخدام معادلة الحساب المكافئة رياضياً للمعادلة (5.14) كما مبين هنا:

$$r=\frac{\sum XY-\frac{(\sum X)(\sum Y)}{N}}{\sqrt{\left[\sum X^2-\frac{(\sum X)^2}{N}\right]\left[\sum Y^2-\frac{(\sum Y)^2}{N}\right]}} \qquad (5.15)$$

حيث:

r = معامل بيرسون

ΣX = مجموع الدرجات في توزيع X

ΣY = مجموع الدرجات في توزيع Y

ΣXY = مجموع حواصل الضرب Y, X الزوجية

ΣX^2 = مجموع مربعات الدرجات في توزيع X

ΣY^2 = مجموع مربعات الدرجات في توزيع Y

N = عدد أزواج الدرجات Y, X الزوجية (الأفراد)

الجدول (5.7): حساب معامل ارتباط بيرسون باستخدام معادلة الدرجات الخام

(1) الأفراد	(2) X	(3) Y	(4) X^2	(5) Y^2	(6) XY
1	18	28	324	784	504
2	18	30	324	900	540
3	17	30	289	900	510
4	17	26	289	676	442
5	16	28	256	784	448
6	16	24	256	576	384
7	15	22	225	484	330
8	15	20	225	400	300
9	14	26	196	676	364
10	14	22	196	484	308
11	13	24	169	576	312
12	13	18	169	324	234
13	12	20	144	400	240
14	12	18	144	324	216
N=14	ΣX=210	ΣY=336	ΣX^2=3206	ΣY^2=8288	ΣXY=5132

وباستخدام بيانات الدرجات الخام نفسها كالسابق، سنستخدم المعادلة (5.15) لحساب معامل ارتباط بيرسون. إن الأعداد والحسابات الضرورية موجودة في الجدول (5.7)، وبتعويض القيم من هذا الجدول بالمعادلة (5.15) يمكن أن نحسب معامل ارتباط بيرسون.

$$r=\frac{5132-\frac{(210)(336)}{14}}{\sqrt{\left[3206-\frac{(210)^2}{14}\right]\left[8288-\frac{(336)^2}{14}\right]}}=+.82$$

لاحظ إنه بتطبيق هذه المعادلة فإنه لم يكن فقط، من غير الضروري تحويل الدرجات إلى درجات Z، بل إن حساب المتوسط الحسابي والانحراف المعياري للتوزيعين قد أستبعد .

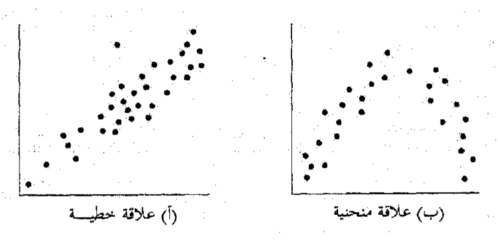
يعود معامل ارتباط بيرسون إلى نفس العائلة الإحصائية مثل المتوسط الحسابي، ويأخذ حسابه بنظر الاعتبار حجم كل درجة في كلا التوزيعين (X) و (Y). وإنه مثل المتوسط الحسابي والانحراف المعياري، يعد إحصاءة فترة كما يمكن استخدامه أيضاً مع البيانات النسبية.

ثمة افتراض يشكل الأساسي لمعامل ارتباط بيرسون، وينص على أن العلاقة بين المتغيرين (X) و (Y) هي علاقة خطية، أي إن هناك خطاً مستقيماً يعطي تعبيراً معقولاً لعلاقة متغير بآخر. وإذا كانت هناك حاجة لخط منحني، للتعبير عن هذه العلاقة فيقال إن العلاقة تشكل خطاً منحنياً (أو علاقة منحنية).

إن الطريقة العملية لإيجاد ما إذا كانت العلاقة بين متغيرين خطية أو خطاً منحنياً، تكون بتفحص بيان التشتت للبيانات. ويبين الشكل (5.9) رسمين بيانيين أحدهما (أ) يوضح علاقة خطية والآخر (ب) يمثل علاقة خط منحني.

وإذا كانت العلاقة بين متغيرين تمثل خطاً منحنياً فإن حساب معامل ارتباط بيرسون سيؤدي إلى سوء تقييم لدرجة العلاقة. وفي هذه الحالة ينبغي تطبيق مؤشر آخر مثل نسبة الارتباط "إيتا" (eta). ويمكن إيجاد النقاش حول نسبة الارتباط في الفصل 11.

الشكل (5.9): علاقات خطية ومنحنية



تفسير معامل ارتباط بيرسون Interpretation of Pearson r

رأينا أنه عند ارتباط متغيرين بشكل عال بطريقة موجبة فإن الارتباط بينهما سوف يقترب من (+1) وعندما يرتبطان بشكل عال بطريقة سالبة فأن الارتباط يقترب إلى (-1) وعندما تكون هناك علاقة ضئيلة بين المتغيرات فالارتباط يكون قريباً من (الصفر). ويقدم معامل ارتباط بيرسون مؤشراً ذا معنى لتوضيح العلاقة، حيث إشارة المعامل توضح اتجاه العلاقة والفرق بين المعامل و (الصفر) توضح درجة العلاقة.

ومع ذلك فإن تفسير معامل الارتباط يقتضي أن يضع المرء النقاط التالية في حسبانه:

1. **لا يوضح الارتباط بالضرورة، السببية.** فإذا تبين أن متغيرين مرتبطان، فإن ذلك يشير إلى أن المواقع النسبية لمتغير ترتبط بالمواقع النسبية للمتغير الآخر. ولا يعني ذلك، بالضرورة، أن التغيرات في متغير واحد تسببها التغيرات في المتغير الآخر. وفي مثالنا، وجدنا ارتباطاً (+0.82) بين درجات اختبار في الإحصاء الوصفي وآخر في الإحصاء الاستدلالي. إن معامل الارتباط هذا يشير إلى أن الشخص الذي ينال درجة فوق المتوسط في اختبار واحد سيحتمل له أن يحصل على درجة فوق المتوسط في الاختبار الآخر .

غير أنه لا يمكننا القول إن التحصيل العالي في اختبار "يسبب" تحصيلاً عالٍ في الاختبار الآخر. فقد تأتي الدرجات في كلا الاختبارين عن أسباب أخرى مثل الاستعداد العددي للأشخاص الذين أخذوا هذه الاختبارات.

وباستخدام مثال آخر، أفترض أننا وجدنا ارتباطاً موجباً بين ثروة عائلات وذكاء أطفال هذه العائلات. ومثل هذا الارتباط لا يدل أبدا على وجود علاقة، سبب بنتيجة، بين هذين المتغيرين. فالثروة لا تؤدي بالضرورة إلى الذكاء ولا يؤدي الذكاء بالضرورة إلى خلق الثروة لهؤلاء الأشخاص.

2. **إن حجم الارتباط هو من ناحية جزئية، دالة التشتت لتوزيعين ستجرى دراسة ترابطهما.** وتقييد مدى الدرجات التي ستجري دراسة ترابطها يقلص من درجة العلاقة الملحوظة بين متغيرين. مثلاً، لوحظ أن النجاح في لعب كرة السلة مرتبط مع الطول: فكلما كان الشخص أطول كان هناك احتمال أكبر في أن يلعب بشكل جيد في هذه الرياضة. وتصح العبارة عن مجموع السكان بشكل عام حيث هناك مدى واسع من الأطوال. وعلى أية حال، ضمن فريق كرة السلة الذي يعتبر كل أعضائه طوال القامة، قد يكون هناك ارتباط قليل أو معدوم بين الطول والنجاح لأن مدى الأطوال مقيد.

وفي كلية تقبل طلبة ذات مدى واسع من الدرجات في اختبار الاستعداد الدراسي، نتوقع ارتباطاً بين درجات الاختبار وتقديرات الكلية. وفي كلية تقبل، فقط، الطلبة الذين درجات استعدادهم عالية جداً، نتوقع ارتباطاً ضئيلاً جداً بين درجات الاختبار والتقديرات بسبب المدى المحدود لدرجات الاختبار في هذا الوضع.

3. **لا ينبغي تفسير معاملات الارتباط بدلالة النسبة المئوية للارتباطات التامة.** فنظراً لأنه يجري التعبير عن معاملات الارتباط كأجزاء عشرية، فإن الأشخاص غير المدربين في الإحصاء يفسرون معاملات الارتباط، أحياناً، على أنها نسبة مئوية لارتباط تام. فمعامل ارتباط يساوي (0.80 = r) لا يفسر 80 في المئة من العلاقة التامة بين متغيرين. فهذا التفسير خاطئ لأن (0.80 لا يدل عن علاقة تكون أكبر مرتين من (0.40 = r). إن إحدى طرق تحديد الدرجة التي يمكن عندها للمرء أن يتنبأ بمتغير من آخر، تكون عن طريق حساب مؤشر يدعى (معامل التحديد). ومعامل التحديد يساوي مربع معامل الارتباط. وهو يعطي النسبة المئوية للتباين، في متغير واحد، الذي يرتبط مع تباين المتغير الآخر. فمثلاً، إذا وجدنا ارتباطاً (0.80) بين التحصيل والذكاء فإن 64 في المئة من التباين في التحصيل ترتبط مع التباين في درجات اختبار الذكاء. وربما كانت أفضل طريقة لإعطاء معنى لحجم معامل الارتباط، هي إعطاء صورة درجة التشتت التي تتضمنها ارتباطات أحجامها مختلفة (كما هو موضح في الشكل 5.8) لتصبح مألوفة مع حجم الارتباطات الملحوظة عموما بين المتغيرات ذات العلاقة.

4. **تجنب تفسير معاملات الارتباط بمعنى مطلق.** فلدى تفسير درجة الارتباط، ضع في الذهن، الهدف الذي يستخدم لأجله. مثلاً، معامل ارتباط (0.50) قد يكون مقنعاً عند التنبؤ بالأداء المستقبلي لمجموعة أشخاص، لكنه قد لا يكون من الحكمة استخدام معامل الارتباط هذا للتنبؤ بأداء شخص واحد في مهمة مستقبلية. فالمعامل (0.50) ليس قيمة مطلقة مع المضمون ذاته في كلا الحالتين.

ويقدم الفصل 11 معلومات إضافية حول استخدام وتفسير الارتباط.

الحاسبات والبرامج الحاسوبية للإحصاء
CALCULATORS AND COMPUTER PROGRAMS FOR STATISTICS

إن لكثير من حاسبات الجيب برامج ذاتية للإجراءات الإحصائية موصوفة في هذا الفصل والفصل التالي. فوصف البرامج الإحصائية المسجلة للحواسيب الكبيرة والصغيرة موجود في الفصل 15.

ما وراء التحليل / تحليل التحليلات META-ANALYSIS

إن "ما وراء التحليل" طريقة لدمج البيانات الكمية من عدد من الدراسات التي تركز على نفس المسألة واستخدام متغيرات مشابهة. ويعرفه (Glass, McCaw, and Smith, 1981) على أنه

"تحليل التحليلات". وباستخدام الإحصائيات المذكورة في عدد من الدراسات الأصلية بمقارنة مجموعات أدخلت عليها معالجة (مجموعات تجريبية) مع مجموعات لم تتعرض لمعالجة (مجموعات ضابطة) ليشق المرء مقياساً عاماً، يُفسر على أنه خلاصة شاملة لنتائج الدراسات المنتخبة.

وبالنسبة لكل دراسة فأن الفرق بين وسط المجموعة التجريبية والضابطة يترجم إلى وحدات انحراف معياري وذلك بتقسيم ذلك الفرق على الانحراف المعياري للمجموعة الضابطة ويدعى هذا المعدل في وحدات الانحراف المعياري "بحجم التأثير". فلكل دراسة:

$$\Delta = \frac{\overline{X}_E - \overline{X}_C}{\sigma_C} \qquad (5.16)$$

حيث:

Δ = حجم التأثير المقدر

$\overline{X}_E$ = المتوسط الحسابي للمجموعة التجريبية

$\overline{X}_C$ = المتوسط الحسابي للمجموعة الضابطة

σ_C = الانحراف المعياري

وقد يفسر كل حجم للتأثير عندئذ، كدرجة (Z) مع متوسط حسابي (صفر) وانحراف معياري هو (1) والمرجع، أو (صفر)، النقطة $\overline{X}_C$ (الوسط الحسابي للمجموعات الضابطة). ويضرب حجم التأثير لكل دراسة بالعدد الكلي للأشخاص في تلك الدراسة، ثم إن مجموع هذه الحواصل يقسم على العدد الكلي للأشخاص في جميع الدراسات للحصول على متوسط حجم التأثير لكل الدراسات. ويفسر هذا المتوسط على أنه أفضل تقدير لاتجاه ومقدار التأثير لمتغير مستقل على متغير تابع.

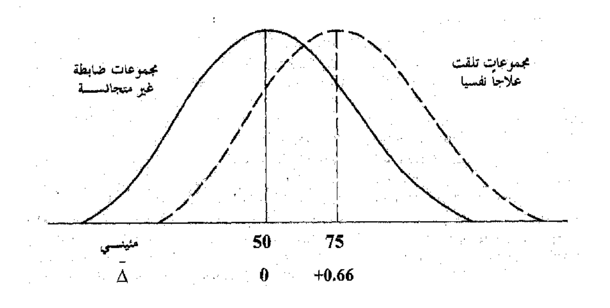
فمثلاً استخدم (Smith & Glass, 977) تحليل التحليلات meta-analysis لاستقصاء المسألة الواسعة: هل يأتي العلاج النفسي بأي اختلاف؟ لقد حدد بحث معياري في الأدبيات 1000 تجربة. ونتج عن التجارب المنتخبة باعتبارها مناسبة للتحليل الكامل، ما مجموعه 833 حجم تأثير. لقد تضمنت الدراسات المنتخبة، الأنا / الذات والحركي والسلوكي واستراتيجيات المعالجة الإنسانية، المرتبطة تجريبياً بالمتغيرات الناتجة مثل احترام الذات والتكيف، قلق الخوف، والأداء المدرسي. لقد كان معدل حجم التأثير هو (0.68)، أي إن معدل المتوسط للمتوسطات الحسابية لما بعد المعالجة للأشخاص المعالجين، كان مساوياً لدرجة (0.68) من الانحراف المعياري فوق المتوسط الحسابي للأشخاص غير المعالجين. واستنتج سميث وكلاس أن النتيجة النموذجية للعلاج النفسي هي زيادة في المتغير التابع، مكافئة لحركة من المتوسط الحسابي إلى المئيني الخامس والسبعين للمجموعة الضابطة. (أنظر الشكل 5.10).

ويستخدم ما وراء التحليل لدمج نتائج الدراسات غير التجريبية. وقام (Hyde, Fennema, & Lamon, 1990) باستخدامه في 100 دراسة تقارن أداء الإناث والذكور في مادة الرياضيات في اختبارات مقننة. ومن مجموع 3,175,188 تلميذاً في هذه المئة دراسة كان المتوسط الحسابي للذكور متواضعاً إذ زاد بمقدار 0.20 من الانحراف المعياري عن المتوسط الحسابي للإناث. وعندما استثنى الباحثون دراسات مثلت عينات منتخبة للأداء المتدني أو الأداء العالي (كطلاب الكلية أو المرتبطين بها) كان متوسط الفرق مجرد (0.05) أعلى لدى الأناث. واستنتج الباحثون أن اختلافات الجنس في أداء الرياضيات كانت صغيرة، بوجه عام.

ويمكن استخدام ما وراء التحليل، ليس فقط، مع دراسات تقارن المتوسطات الحسابية، بل أيضاً مع دراسات الارتباط والنسبة وغيرها من المقاييس. ولمزيد فيما يتعلق بما وراء التحليل وأشكال أخرى من الطرق الكمية لتوليف البحوث، هناك مرجع مفيد هو:

The Handbook of Research Synthesis (cooper & Hedges, 1993)

الشكل (5.10): معدل حجم التأثير التقديري للعلاجات، مقارنة مع مجموعات غير معالجة (البيانات مستمدة من 40,000 شخص، بعضهم خضع للعلاج وآخرون لم يعالجوا وبمعدل أقل من 20 ساعة علاج مع معالجين متوسط خبراتهم 2½ سنة)



الخلاصة SUMMARY

تفيد الإحصائيات الوصيفة في وصف وتلخيص الملاحظات. فالأسلوب الوصفي المستخدم هو أسلوب يجري انتقاؤه وفقاً لهدف دور الإحصاءة والمقياس المستخدم في القياسات. أما معايير القياس فهي وسائل قياس الملاحظات وهي على أربعة أنواع: 1) المقاييس الأسمية التي تصنف الملاحظات في أصناف منفصلة تماماً، 2) المقاييس الترتيبية التي تفرز الأشياء أو الأصناف على أساس مرتبتها النسبية، 3) مقاييس الفترة التي تستخدم فترات متساوية للقياس وتوضح درجة امتلاك الشخص أو الشيء لصفة معينة، 4) مقاييس نسبة، تستخدم الفترات المتساوية للقياس وهي تقيس من نقطة صفر مطلقة.

وحالما يتم التعبير عن الملاحظات، يمكن ترتيب البيانات في توزيعات تكرارية وإيضاحها بيانياً في مضلعات أو مدرجات تكرارية.

إن قياسات النزعة المركزي- المنوال، الوسيط والمتوسط / الوسط الحسابي- تؤمن مؤشراً واحداً لتمثيل القيمة المتوسطة لجملة القياسات كلها.

إن المنوال ، الذي يعد إحصاءة أسمية هو القياس الأقل استقراراً والأقل فائدة في بحوث التربية. أما الوسيط فهو إحصاءة ترتيبية آخذين بعين الاعتبار مراتب الدرجات ضمن توزيع معين وليس حجم الدرجات الفردية. أما المتوسط الحسابي، وهو إحصاءة فترية (أو نسبية) يعد المؤشر الأكثر استقرار والأوسع استخداماً للنزعة المركزية.

ويمكن بطريقة أخرى وصف الملاحظات، بالإشارة إلى تغير القيم أو انتشارها ضمن التوزيع. أما المدى والانحراف الربيعي والتباين والانحراف المعياري فهي أربعة مؤشرات تستخدم لهذا الغرض. والمدى، إحصاءة أسمية وهو المسافة بين أعلى وأدنى القيم في توزيع مضافاً لها واحد. ويعطي الانحراف الربيعي نصف المسافة بين أعلى وأدنى الأرباع/ الربيعات. وهو إحصاءة ترتيبية، أما التباين فهو المتوسط الحسابي لمربعات انحرافات القيم عن المتوسط الحسابي، وهو إحصاءة فترية أو نسبية، بينما الانحراف المعياري وهو الجذر التربيعي للتباين فهو المؤشر الأوسع استخداماً للتغير/ التشتت (الانتشار).

تستخدم الدرجات المعيارية لإيضاح موقع درجة واحدة في توزيع معين. والأوسع استخداماً في هذا المجال هي الدرجة (z) التي تحول القيم إلى وحدات انحراف معياري. وباستخدام صفات ومناطق المنحنى العادي يمكن أن نقرب النسبة المئوية للحالات دون وفوق كل درجة (z) في التوزيع العادي.

وتمكننا أساليب الارتباط من وصف العلاقة بين جملتين من القياسات. ومعامل ارتباط بيرسون (r) هو الأوسع استخداماً كمؤشر للعلاقة. ويستخدم هذا المعامل مع بيانات الفترات أو

النسبة. يلخص الجدول 5.8 هذه الإحصاءات (*)(مجموع إحصاءة).

الجدول (5.8): ملخص الإحصاءات الوصفية المعروضة في هذا الفصل

	أسمي	ترتيبي	فترة
مؤشرات النزعة المركزية	منوال	وسيط	متوسط حسابي
مؤشرات المتغيرية/ التشتت	مدى	انحراف ربيعي	التباين والانحراف المعياري
مؤشرات الموقع	تصنيف	مرتبة ميئنية/ مئوية	درجة z ، درجة T ، ودرجات معيارية أخرى
مؤشرات ارتباط	فاي *(Phi)	سبيرمان (Rho)	معامل بيرسون (r)

* إن معاملات ارتباط (فاي) وسبيرمان موصوفة في الفصل 11.

ويدمج "ما وراء التحليل- أو تحليل التحليلات" البيانات من دراسات عدة في مؤشر واحد، يدعى، حجم التأثير. ويوضح هذا المؤشر متوسط الفرق بين المتوسطات الحسابية للمجموعتين التجريبية والضابطة كما قيست بالانحراف المعياري للمجموعة الضابطة لكل دراسة.

مفاهيم أساسية Key Concepts

coefficient of determination	معامل التحديد
correlation	ارتباط
correlation coefficient	معامل ارتباط
cumulative frequency	تكرار تراكمي
curvilinear relationship	علاقة خطية منحنية أو علاقة منحية، اختصاراً
descriptive statistics	إحصاء وصفي
effect size	حجم التأثير
frequency distrihution	توزيع تكراري
frequency polygon	مضلع تكراري
histogram	مدرج تكراري
inferential statistics	إحصاء استدلالي
interval scale	مقياس فترة

(*) تشير هنا إلى مجموع إحصاءة (Statistic) التي تحدد صفة أو خاصة للتوزيع (المراجع).

linear relationship	علاقة خطية
mean	المتوسط الحسابي (أو الوسط الحسابي أو الوسط اختصاراً)
median	الوسيط
measures of central tendency	مقاييس النزعة المركزية
meta-analysis	ما وراء التحليل أو تحليل التحليلات
mode	المنوال
negative correlation	الارتباط السالب
negatively skewed curve	المنحنى المنحرف/ الملتوي سلبيا
nominal scale	مقياس اسمي
normal distribution	توزيع اعتيادي/ عادي / نظامي
ordinal scale	مقياس ترتيبي
Pearson product moment coefficient	معامل بيرسون للارتباط التتابعي
positive correlation	ارتباط موجب
positively skewed curve	منحنى منحرف/ ملتو إيجابياً
quartile deviation (semiinterquartile range)	انحراف ربيعي/ إرباعي (مدى شبه إرباعي)
range	مدى
ratio scale	مقياس نسبي
scattergram	بيان التشتت/ الانتشار
skewed distribution	توزيع منحرف/ ملتوي
standard deviation	انحراف معياري
standard score	درجة معيارية
symmetric distribution	توزيع متناظر/ متماثل
T-score	الدرجة تي (T)
variability	المتغيرية/ التغير/ الانتشار/ التشتت
variance	التباين
z-score	الدرجة (z)

تماريــن EXERCISES

1. حدد نمط ميزان القياس – اسمي ، ترتيبي، فتري أو نسبي – الذي تشير إليه كل عبارة.
   أ. أنهى جون اختبار الرياضيات في 35 دقيقة بينما أنهى جاك نفس الامتحان في 25 دقيقة.
   ب. جاك يتحدث الفرنسية بينما لا يتحدثها جون.
   ج. جاك أطول من جون.
   د. جون طوله (6) أقدام و (2) بوصة.
   هـ. معامل ذكاء جون هو 120 بينما معامل ذكاء جاك هو (110).

2. ارسم مدرج تكرار ومضلع تكرار للتوزيع التكراري التالي:

X	*f*	*X*	*f*	*X*	*f*	*X*	*f*
80	1	76	6	73	20	70	7
79	2	75	15	72	17	69	3
78	3	74	22	71	9	68	1
77	10						

3. أعط إجابة كما هو مطلوب، إذا أخذنا بنظر الاعتبار التوزيع التالي:
   15, 14, 14, 13, 11, 10, 10, 10, 8, 5
   أ. أحسب الوسط الحسابي.
   ب. حدد قيمة الوسيط.
   ج. حدد قيمة المنوال.

4. فسر باختصار العلاقة بين انحراف/ التواء توزيع درجات معينة والقيم الناتجة للمتوسط الحسابي، الوسيط، المنوال.

5. حدد المقياس – المنوال، الوسيط، أو المتوسط الحسابي– الذي يناسب كل نوع من الموازين:
   أ. ترتيبي ب. اسمي ج. فاصلي

6. حدد المقياس – منوال، وسيط أو متوسط حسابي– الذي يحدده كل مصطلح:
   أ. درجة الوسط ب. المتوسط الحسابي ج. الدرجة الأكثر حدوثاً.

7. تمثل الدرجات أدناه درجات اختبار المفردات لعشرين تلميذاً في الصف السابع. أحسب المدى والانحراف المعياري والانحراف الربيعي، وناقش الفوائد والمساوئ لكل منها كمقياس لتغيير الدرجات.

X	f	fX	X^2	fX^2
16	1	16	256	256
15	0	0	225	0
14	0	0	196	0
13	0	0	169	0
12	2	24	144	288
11	0	0	121	0
10	2	20	10	200
9	1	9	81	81
8	1	8	64	64
7	1	7	49	49
6	4	24	36	144
5	2	10	25	50
4	1	4	16	16
3	1	3	9	9
2	4	8	4	16

8. لتقليص أثر درجة متطرفة: هل ينبغي للمرء أن يختار الانحراف الربيعي أم الانحراف المعياري كمقياس للتغير؟ لماذا؟

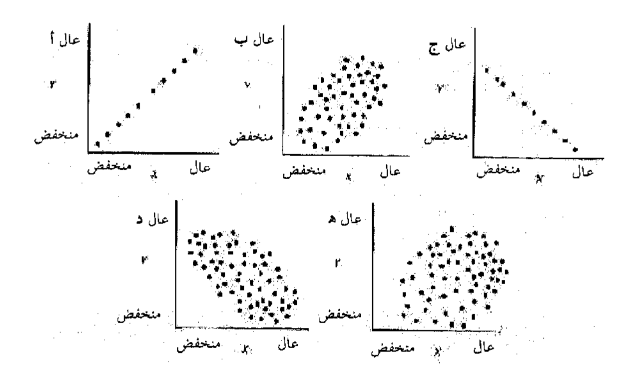
9. درجة الوسط الحسابي في اختبار هي (40) والانحراف المعياري هو (4). عبر عن كل من الدرجات الخام التالية كدرجات (z).

أ. (41) ب. (30) جـ. (48) د. (36) هـ. (46)

10. ما هي الدرجة T للدرجة 46 الخام في التمرين (9).

11. في توزيع اعتيادي، ما هي النسبة المئوية للدرجات التي تقع تحت $z = -1$ ؟ ونقطة $z = 0$ ؟ و$z = +0.67$ ؟.

12. صف العلاقة المبينة ببيان التشتت في الأشكال التالية، ثم قدر معاملات الارتباط.



13. أخذ خمسة تلاميذ اختبار التاريخ واختبار الجغرافيا وحصلوا على النتائج التالية:

	تاريخ		جغرافيا	
	درجة خام	درجة Z	درجة خام	درجة Z
جورج	28	.5	85	1.5
رالف	32	1.5	65	.5
ديفيد	26	0	55	0
بول	20	-1.5	45	-.5
جون	14	-.5	25	-1.5

تاريخ	جغرافيا
$\Sigma X = 130$	$\Sigma X = 275$
$\sigma = 4$	$\sigma = 20$

أ. ما هو الوسط الحسابي لاختبار التاريخ؟

ب. من الذي تحصيله في التاريخ هو الأكثر توافقاً مع تحصيله في الجغرافيا؟

جـ. ما هو الارتباط بين درجات التاريخ والجغرافيا؟

14. إذا كان المتوسط الحسابي لاختبار التاريخ في التمرين 13 أدنى من المتوسط الحسابي لاختبار الجغرافيا، فأي من الاستنتاجات التالية سيكون واضحاً.

أ. هؤلاء الصبيان أفضل في التاريخ منه في الجغرافيا.
ب. هؤلاء الصبيان أفضل في الجغرافيا منه في التاريخ.
جـ. ربما أنفق مدرسهم وقتاً على الجغرافيا أكثر مرتين منه على التاريخ.
د. يعرف مدرسهم الجغرافيا أكثر من التاريخ.
هـ. ليس أي مما ورد أعلاه.

15. إذا كان معامل الارتباط بين المتغير (X) والمتغير (Y) هو (0.98-) فأي مما يلي سيتم إيضاحه؟
أ. المتغير (X) والمتغير (Y) مرتبطان بشكل دقيق.
ب. المتغير (X) والمتغير (Y) غير مرتبطين.
جـ. المتغير (X) والمتغير (Y) مرتبطان بشكل تام.
د. المتغير (Y) نتيجة للمتغير (X).

16. بالنسبة لكل من الحالات التالية، حدد الإحصاءة التي ينبغي استخدامها – المتوسط الحسابي، الانحراف الربيعي، الدرجة (Z)، أو معامل ارتباط بيرسون (r).
أ. نريد أن نعرف مدى انتشار أو عدم تجانس درجات صف معين.
ب. نريد أن نحدد كيف نقارن درجة (جون) مع درجات بقية الصف.
جـ. نريد أن نعرف مدى حسن أداء الصف، ككل، في امتحان معين.
د. نريد أن نتنبأ بالتحصيل المستقبلي للطلبة من معامل ذكائهم.

17. ما الأسلوب الإحصائي الذي يدعى بتحليل التحليلات؟

الأجوبة ANSWERS

1. أ. نسبي ب. اسمي جـ. ترتيبي د. نسبي هـ. فتري.
2. ستختلف الإجابات.
3. أ. المتوسط الحسابي=11 ب. الوسيط = 10.5 جـ. المنوال = 10.
4. المقاييس الثلاثة ليست متساوية في التوزيع المنحرف، فالمتوسط الحسابي ينجذب في اتجاه الجانب المنحرف. وهكذا، ففي التوزيع المنحرف إيجابياً، يكون المتوسط الحسابي دائماً أعلى من الوسيط وأن المنوال هو عادة الأدنى في القيمة.

وفي التوزيع المنحرف سلبياً يكون المتوسط الحسابي دائماً أدنى من الوسيط والمنوال هو الأعلى عادة في القيمة.

5. أ. الوسيط ب. المنوال جـ. الوسط الحسابي

6. أ. الوسيط ب. الوسط الحسابي جـ. المنوال

7. المدى = 15 = 1 + (16-2).

$$\sigma=\sqrt{\frac{1173-\frac{(133)^2}{20}}{20}}=\sqrt{14.4275}=3.798$$

$$QD=\frac{Q2-Q1}{2}=\frac{9.5-3.5}{2}=\frac{6}{2}=3$$

8. الانحراف الربيعي لأن الدرجات المتطرفة لا تؤثر على الانحراف الربيعي.

9. أ. 25 ب. -2.5 جـ. 2 د. -1 هـ. 1.5

10. T = 10z + 50 = 10 (1.5) + 50 = 65

11. 16%, 50%, 75%

12. أ. موجب تام، 1+
ب. موجب، 75.+
جـ. سالب تام، 1−
د. سالب، 75.−
هـ. لا يوجد ارتباط، 0

13. أ. 26
ب. ديفيد
جـ. $r=\frac{\sum z_x Z_y}{N}=\frac{3}{5}=.6$

14. هـ

15. أ

16. أ. انحراف ربيعي ب. درجة Z جـ. وسط د. معامل بيرسون

17. ما وراء التحليل.

REFERENCES المصادر

Cooper, H., and Hedges, L.V. (1993). The handbook of research synthesis. New York: Russell Sage Foundation.

Glass, G.V, McGaw, B., and Smith, M.L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage Publications.

Hyde, J.S., Fennema, E., and Lamon, S.J. (1990). Gender differences in mathematics performance: A meta-analysis. Psychological Bulletin, 107, 139-155.

Smith, M.L., and Glass, G.V (1977). Meta-analysis of psychotherapy outcome studies. American Psychologist, 32, 752-760.

Stevens, S.S. (1951). Mathematics, measurement, and psychophysics. In S.S. Stevens (Ed.), Handbook of experimental psychology (p. 1). New York: John Wiley.



المعاينة والاستدلال الإحصائي

Sampling and Inferential Statistics

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون بوسع الطالب أن:

1- يصف المعنى والأساس المنطقي والخطوات التي تنطوي عليها المعاينة والتمييز بين المعاينة الاحتمالية وغير الاحتمالية.

2- يدرج الخواص والاستخدامات والقيود لكل نوع من أنواع المعاينة الاحتمالية واللااحتمالية.

3- يشرح معنى خطأ المعاينة وعلاقته بصنع الاستدلالات الإحصائية.

4- يشرح معنى الدلالة الإحصائية.

5- يفسر معنى الفرضية الصفرية واستخدامها في البحث العلمي.

6- يصف أخطاء النوع 1 والنوع 2.

7- يشرح الفرق بين اختبارات الدلالة الإحصائية الاتجاهية واللا اتجاهية والاستخدام المناسب لكل منها.

8- يصف العناصر التي ينبغي اعتبارها في تحديد حجم العينة.

9- يستخدم معادلات / قوانين القوة لحساب حجم العينة المطلوبة في اختبار إحصائي حول المتوسط الحسابي لمجتمع إحصائي .

10- يشرح الاستراتيجية المستخدمة في اختبار الفرضية الصفرية.
11- يطبق اختبار (t) لإيجاد دلالة الفرق بين الأوساط المترابطة وغير المترابطة.
12- يصف هدف اختبار (t) لمعامل ارتباط بيرسون (r) وتحديد ما إذا كان أي معامل ارتباط معطى، يختلف عن الصفر بدلالة إحصائية.
13- يطبق اختبار F لإيجاد دلالة الفروقات بين المجموعات في تحليل التباين ذي الطريق الواحد والطريقين.
14- يطبق اختبار مربع كاي لإيجاد دلالة الفروقات بين النسب في التصنيفات ذات الطريق/ الاتجاه الواحد والطريقين/ الاتجاهين.
15- يختار نوع الإجراءات الإحصائية الاستدلالية المناسبة للاستخدام في اختبار فرضية بحث معينة.
16- يبين استيعابه للمصطلحات الإحصائية الفنية الأساسية المستخدمة في تقرير نتائج البحث.

تستخدم الإحصائيات التي نوقشت في الفصل السابق لتنظيم وتلخيص ووصف البيانات. وعلى أية حال، غالباً ما نحتاج في البحوث إلى الذهاب أبعد من مجرد وصف البيانات. فبعد إجراء ملاحظات عينة ما، نستخدم الاستقراء أو الاستدلال لتعميم استنتاجاتنا على كل المجتمع الإحصائي الذي استقيت منه العينة. وللقيام بذلك، نحتاج إلى أساليب تمكننا من إجراء استدلالات صحيحة من عينات إلى كل المجتمعات الإحصائية.

المعاينة SAMPLING (*)

إن إحدى الخواص المهمة للإحصاء الاستدلالي هي عملية السير من الجزء إلى الكل. فمثلاً، قد ندرس مجموعة منتخبة بصورة عشوائية من 500 تلميذ يلتحقون بجامعة بغية إجراء تعميمات حول جماعة الطلاب في تلك الجامعة.

تدعى المجموعة الصغيرة التي لوحظت "بالعينة"، والمجموعة الكبيرة التي جرى عليها التعميم تدعى "المجتمع الإحصائي". ويعرف المجتمع الإحصائي على أنه جميع أفراد أي طبقة محددة من الناس أو الأحداث أو الأشياء. فمثلاً، في الدراسات التي يكون المراهقون الأمريكان فيها المجتمع الإحصائي المعني، يمكن للمرء أن يعرّف هذا المجتمع على أنه جميع البنين والبنات الأمريكان داخل المدى العمري 12-21 سنة. أما العينة فهي جزء من المجتمع. فمثلاً، طلبة ثانوية واشنطن

(*) المعاينة (Sampling): مصطلح يشير إلى إجراءات اختيار العينة التي يفترض أن تكون ممثلة للمجتمع الإحصائي، بهدف التوصل إلى تعميمات حول هذا المجتمع - (المراجع) .

في أنديانا بولس، تشكل عينة من البالغين الأمريكان. وهم جزء من مجتمع أكبر من حيث كونهم مواطنين أمريكيين، وضمن المدى العمري 12-21 سنة.

أما الاستدلال الإحصائي فهو إجراء يمكن بواسطته تقدير معالم / بارمترات (جمع معلم Parameter أي خاصة/ سمة للمجتمع) من إحصاءات (جمع إحصاءة Statistic أي خاصة/ سمة للعينة). إن مثل هذه التقديرات قائمة على قوانين الاحتمال، وهي أفضل التقديرات وليست حقائق مطلقة. وتتضمن أي من هذه الاستدلالات درجة معينة من الخطأ. وكما سنرى لاحقاً فإن الإحصاء الاستدلالي يستخدم أيضاً لاختبار فرضيات حول مجتمعات إحصائية على أساس ملاحظات العينة المستمدة من المجتمع الإحصائي.

الأساس المنطقي للمعاينة Rationale of Sampling

يعد التفكير الاستقرائي جزءاً أساسياً من الطريقة العلمية. وتنطوي الطريقة الاستقرائية على أجراء ملاحظات والتوصل إلى نتائج من هذه الملاحظات. فإذا كان بوسع المرء ملاحظة جميع حالات المجتمع الإحصائي فسيكون واثقاً بقدرته على التوصل إلى نتائج خاصة بالمجتمع الإحصائي من هذه الملاحظات (استقراء تام). ومن جهة أخرى، إذا لاحظ أحدهم بعض حالات المجتمع فقط، فليس بوسعه أن يقوم بأكثر من الاستنتاج بأن هذه الملاحظات ستكون صحيحة للمجتمع ككل (استقراء غير تام). وهذا هو مفهوم المعاينة الذي ينطوي على أخذ جزء من المجتمع الإحصائي، وإجراء ملاحظات على هذه المجموعة الصغيرة، ثم تعميم الاستنتاجات على المجتمع الإحصائي الأصلي، وهو المجتمع الأكبر الذي استمدت العينة منه.

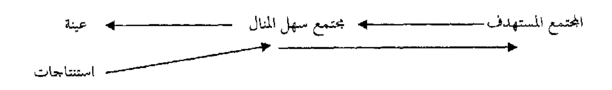
إن المعاينة أمر لا مناص منه بالنسبة للباحث. وبصورة اعتيادية، فإن ما يلزم من وقت ومال وجهد لا يتيح للباحث دراسة جميع أفراد المجتمع الإحصائي المحتملين. علاوة على ذلك، ليس من الضروري، عموماً، دراسة جميع الحالات المحتملة لفهم الظاهرة قيد الدرس. وتنجدنا المعاينة بتمكيننا دراسة جزء من المجتمع بدلاً من المجتمع كله. ولأن هدف سحب العينة من المجتمع الإحصائي هو الحصول على معلومات تخص ذلك المجتمع، فإن من المهم جدا أن يشكل الأفراد المشمولين في العينة مقطعاً عرضياً ممثلاً للأفراد في المجتمع الإحصائي .ويجب أن تكون العينات ممثلة إن أراد المرء أن يكون قادراً على التعميم بثقة من العينة إلى المجتمع. مثلاً، قد يفترض الباحث أن الطلبة في ثانوية واشنطن يمثلون المراهقين في أمريكا، ومع ذلك قد لا تكون هذه العينة ممثلة إذا كان للأفراد المعنيين بعض الخواص التي تختلف عن المجتمع الإحصائي الأصلي. فموقع مدرستهم وخلفياتهم الاجتماعية الاقتصادية وأوضاعهم العائلية وتجاربهم السابقة والعديد من الخواص الأخرى لهذه المجموعة، قد تجعلهم لا يمثلون المراهقين الأمريكيين. فهذا النمط من العينة يدعى "بالعينة المتحيزة". ولا يمكن لاستنتاجات أية عينة متحيزة في دراسة بحثية، أن تعمم بشكل مشروع على المجتمع الذي أخذت منه.

خطوات في المعاينة Steps in Sampling

إن أول خطوة في المعاينة هي تحديد المجتمع الإحصائي الممثل في الدراسة. فإذا كان الباحث مهتماً بمعرفة المدرسين في النظام المدرسي العام في سانت لويس، فإن جميع الذين يقومون بالتدريس داخل ذلك النظام يكونون المجتمع المستهدف. وفي دراسة لمواقف وقيم المراهقين من الأمريكيين، سيكون المجتمع المستهدف هو جميع الفتيات والصبيان الأمريكيين في المدى العمري 12–21 سنة، بفرض أن سن المراهقة يحدد إجرائياً بالفترة ما بين 12 و 21 عاماً.

وعلى أية حال، طالما أنه ليس من الممكن التعامل مع كل المجتمع المستهدف، فيجب على المرء أن يحدد ذلك الجزء من المجتمع الذي يمكن الوصول إليه – ويدعى هذا بالمجتمع السهل المنال – ومن هذه المجموعة يأخذ الباحث عينة للدراسات. وتتأثر طبيعة المجتمع سهل المنال، بوقت ومصادر الباحث. ففي دراسة نموذجية للاتجاهات، مثلاً، قد يخصص الباحث جميع البالغين من صبيان وبنات في كاليفورنيا أو أولئك الذين في سان فرانسيسكو كمجتمع سهل المنال.

وينتخب الباحث من المجتمع سهل المنال، عينة بطريقة تمثل ذلك المجتمع. مثلاً، على الباحث أن يأخذ عينة من المراهقين من جميع ولاية كاليفورنيا، إن تم تحديد المراهقين في كاليفورنيا كمجتمع سهل المنال. أو إن كان البالغون الذين يعيشون في سان فرانسيسكو هم المجتمع السهل المنال، فإن العينة سوف تستمد من تلك الجماعة المعينة.



لكن ما مدى الأمان في التعميم من العينة إلى المجتمع المستهدف؟ فإذا كانت العينة المنتخبة ممثلة فعلاً للمجتمع المستهدف، فإن هناك صعوبة قليلة في إجراء الخطوة الأولى في عملية التعميم. والقاعدة العامة هي: إذا انتخبت عينة بحث تمثل مجتمعاً يمكن الوصول إليه، فإنه يمكن تعميم الاستنتاجات من العينة على ذلك المجتمع. فمثلاً إذا انتخب أحدهم عينة ممثلة من المراهقين في كاليفورنيا فإن بوسعه إجراء تعميمات حول اتجاهات وقيم جميع المراهقين من الصبيان والبنات في كاليفورنيا.

والتعميم، على أية حال، من المجتمع الإحصائي الذي يمكن الوصول إليه، إلى المجتمع المستهدف ينطوي على مجازفة أكبر. فالثقة التي لدى المرء في هذه الخطوة، تعتمد على التشابه بين المجتمع الممكن الوصول إليه، والمجتمع المستهدف. ففي المثال أعلاه، ربما كان للباحث ثقة أكبر في إقامة التعميمات حول المراهقين الأمريكيين لو أن المراهقين في ولايات متعددة عبر البلد

قد تم تصنيفهم كمجتمع سهل المنال، وليس هؤلاء الذين في كاليفورنيا وحدها. بهذه الطريقة ستتمثل جميع شرائح الولايات المتحدة وسيكون بالإمكان أخذ عينة مناسبة للاتجاهات والقيم.

صحيح، أن على المرء أن يقوم "بقفزة ثقة" استدلالية عند تقرير خواص المجتمع من ملاحظات العينة. واحتمال صحة مثل هذه الاستدلالات يعتمد بشكل كبير على الإجراءات المستخدمة في المعاينة.

وهناك نوعان أساسيان من إجراءات المعاينة متيسران للباحثين: معاينة احتمالية ومعاينة غير احتمالية. فالمعاينة الاحتمالية تنطوي على انتخاب عينة يتم اختيار عناصرها عن طريق إجراءات الصدفة. والسمة الرئيسية للمعاينة الاحتمالية هي أن لكل فرد أو عنصر من المجتمع الإحصائي، احتمال معروف لاختياره في العينة. أما المعاينة الاحتمالية فتشمل طرق اختيار لا يتم فيها اختيار العناصر عن طريق إجراءات الصدفة. وليس هناك من طريقة لتقدير احتمال اشتمال كل عنصر في العينة. ويعتمد بحاجة على المعرفة والخبرة وحكم الباحث. وتستخدم المعاينة غير الاحتمالية عندما يكون تطبيق المعاينة الاحتمالية غير عملي. ومزاياها هي الملاءمة والاقتصاد.

المعاينة الاحتمالية (معاينة الاحتمال) Probability Sampling

تعرف المعاينة الاحتمالية على أنها نوع من المعاينة يكون لكل عنصر فيها من المجتمع فرصة لا صفرية في الانتخاب (Stuart, 1984). إن التضمين المحتمل لكل عنصر في المجتمع الإحصائي في هذا النوع من العينة يحدث بالصدفة، ويتم الحصول عليه من خلال الانتخاب العشوائي. وعند استخدام معاينة الاحتمال فإن الإحصاء الاستدلالي يتيح للباحثين تقدير مدى احتمال اختلاف، الاستنتاجات القائمة على العينة، عن تلك التي قد يجدها الباحث فيما لو تمت دراسة المجتمع الإحصائي كله. إن الأنواع الأربعة من معاينة الاحتمال المستخدمة عادة في بحوث التربية هي المعاينة العشوائية البسيطة، والمعاينة الطبقية، والمعاينة العنقودية، والمعاينة المنتظمة.

المعاينة العشوائية البسيطة Simple Random sampling

إن أفضل إجراءات معاينة الاحتمال هي المعاينة العشوائية البسيطة. والخاصة الرئيسية لهذا النوع هي أن جميع أفراد المجتمع لديهم فرصة متساوية ومستقلة في الانضمام إلى العينة. وبعبارة أخرى، ففي كل زوج من العناصر (X) و (Y) يكون فرصة (X) في الانتخاب مساوية لفرصة (Y)، وأن انتخاب (X) لا يؤثر بأية طريقة على احتمال انتخاب (Y)، وتضم خطوات المعاينة العشوائية البسيطة ما يلي:

1- تعريف المجتمع الإحصائي.
2- إدراج كل أفراد المجتمع الإحصائي في قائمة.
3- انتخاب العينة باستخدام أجراء، تقرر فيه الصدفة فقط أعضاء القائمة الذين سينضمون للعينة.

والأجراء الأكثر انتظاماً في المعاينة العشوائية تكون بالعودة إلى جدول "الأعداد العشوائية". وهو عبارة عن جدول يحتوي على أعمدة بالأرقام المولدة بشكل ميكانيكي، باستخدام الحاسوب عادة لضمان النسق العشوائي، والجدول 6-A في الملحق مثال على ذلك. إن الخطوة الأولى في تكوين عينة عشوائية من مجتمع ما، هي تخصيص عدد محدد لكل عضو في المجتمع الإحصائي. ثم يتم استخدام جدول الأعداد العشوائية لانتخاب الأعداد المحددة للأفراد المشمولين في العينة.

دعنا نوضح استخدام هذا الجدول للحصول على عينة من المراهقين من الطلبة الملتحقين بثانوية واشنطن. أولاً من الضروري عدّ جميع أفراد المجتمع. ويمكن لمكتب المدير أن يقدم قائمة بالطلبة المسجلين في المدرسة. ثم يخصص المرء عدداً لكل فرد في المجتمع لأغراض التعريف. وعلى افتراض أن هناك 800 طالب في المدرسة، يخصص المرء عدداً لكل فرد في المجتمع لأغراض التعريف. وحسب افتراض أن هناك 800 طالب في المدرسة فإن المرء قد يستخدم الأعداد ,000) (799 ,... ,001, 002, 003 لهذه الغاية، ثم يقوم بإدخال جدول الأعداد العشوائية للحصول على أعداد كل منها من ثلاثة أرقام وذلك فقط باستخدام الأعداد التي هي أقل من (799) أو مساوية لها. وبالنسبة لكل عدد مختار فإن العنصر المطابق له في المجتمع، يقع داخل العينة. ويستمر المرء بالعملية حتى يتم اختيار العدد المطلوب للعينة. ومن المعتاد استخدام جدول الأعداد العشوائية لتحديد النقطة —بطريق الصدفة – التي عندها يتم دخول الجدول. وإحدى الطرق هي لمس الصفحة بشكل أعمى والبدء حيثما لمست الصفحة.

إن المعنى المفهوم لكلمة (عشوائي) هو "بدون هدف أو بالمصادفة". ومع ذلك فإن المعاينة العشوائية هي هادفة ومنهجية. فمن الواضح أن العينة المنتجة عشوائياً ليست عرضة لتحيزات الباحث. بل إن الباحثين يلزمون أنفسهم بانتخاب عينة بطريقة لا تسمح لتحيزاتهم بالتدخل. فهم يعدون أنفسهم بتجنب الانتخاب المبيت للأشخاص الذين سيؤكدون الفرضية. أنهم يتيحون للصدفة وحدها أن تحدد العناصر في المجتمع الإحصائي التي ستكون في العينة.

قد يتوقع المرء أن تكون العينة العشوائية ممثلة للمجتمع الأصلي المُعاين. وعلى أية حال، فإن الانتخاب العشوائي وخصوصاً في العينات الصغيرة، لا يضمن بصورة مطلقة عينة تمثل المجتمع بصورة جيدة. ولا يضمن الانتخاب العشوائي أن تكون الفروقات بين العينة والمجتمع الأصلي هي من باب الصدفة فقط، وليس نتيجة تحيز الباحث. والفروقات بين العينات العشوائية ومجتمعها الأصلي ليست نظامية. فمثلاً، قد يكون الوسط الحسابي للتحصيل القرائي في عينة عشوائية لطلبة الصف السادس، أعلى من الوسط الحسابي للتحصيل القرائي للمجتمع الأصلي، إلا أنه من المحتمل أيضا أن يكون وسط العينة أدنى من وسط المجتمع الإحصائي الأصلي. بعبارة أخرى، فإن احتمال أن تكون أخطاء المعاينة في المعاينة العشوائية سالبة كما هي موجبة. وعلاوة على ذلك، أوضح علماء النظريات الإحصائية من خلال التفكير الاستنتاجي مقدار التباين بين ما يتوقعه المرء من ملاحظات مستمدة من عينات عشوائية، وما لوحظ في المجتمع الإحصائي.

فجميع إجراءات الإحصاء الاستدلالي تضع هذا الهدف في البال. تذكر إن احتمال التباين في الخواص الملاحظة في عينة صغيرة عن خواص المجتمع الإحصائي أكثر من الخواص الملاحظة في عينة كبيرة. فعند استخدام معاينة عشوائية يمكن للباحث أن يستخدم الإحصاء الاستدلالي لتقرير مقدار احتمال تباين المجتمع الإحصائي عن العينة. ويستند الإحصاء الاستدلالي في هذا الفصل دائماً على المعاينة العشوائية وينطبق مباشرة فقط، على تلك الحالات التي تكون فيها المعاينة عشوائية.

ولسوء الحظ، تتطلب المعاينة العشوائية البسيطة تعداد جميع الأشخاص في مجتمع إحصائي منته(*) قبل سحب العينة – وهذا شرط غالباً ما يطرح عائقاً أمام استخدام هذه الطريقة بصورة عملية.

المعاينة الطبقية Stratified Sampling

عندما يتكون المجتمع الإحصائي من عدد من المجموعات الجزئية أو الطبقات التي قد تختلف في الخواص المدروسة، فإنه من المرغوب فيه، غالباً، استخدام شكل من المعاينة الاحتمالية يدعى (المعاينة الطبقية). فمثلاً، إذا كان أحدهم يجري استفتاءاً مصمماً لتقييم الآراء حول قضية سياسية معينة، فقد يكون من المناسب تقسيم المجتمع الإحصائي إلى مجموعات جزئية على أساس العمر، أو المهنة، لأن المرء قد يتوقع أن تكون الآراء مختلفة بصورة منتظمة بين المجموعات العمرية أو المهنية المتنوعة. وفي المعاينة الطبقية يحدد المرء أولا الطبقات المعنية، ثم يسحب بصورة عشوائية عدداً محدداً من الأشخاص من كل طبقة. وقد يكون الأساس التصنيفي (الطبقي) جغرافياً أو قد ينطوي على خواص للمجتمع الإحصائي، كالدخل، أو المهنة، أو الجنس، أو العمر، أو السنة في الكلية، أو المستوى التعليمي. ففي دراسة المراهقين مثلاً، قد يهتم المرء ليس فقط في مسح اتجاهات المراهقين نحو ظواهر معينة، بل كذلك في مقارنة اتجاهات المراهقين الذين يقيمون في مدن صغيرة مع من يعيشون في مدن متوسطة أو كبيرة الحجم. ففي مثل هذه الحالة يقسم المرء المجتمع الإحصائي للمراهقين إلى ثلاث مجموعات استناداً إلى حجم المدن التي يعيشون فيها، ثم ينتخب عشوائياً عينات مستقلة من كل طبقة . وتكمن مزية المعاينة الطبقية في كون أنها تمكن الباحث من دراسة الفروقات التي قد تكون موجودة بين المجموعات الفرعية الجزئية المتنوعة من المجتمع الإحصائي. ففي هذا النوع من المعاينة قد يأخذ المرء أعداداً متساوية من كل طبقة أو ينتخب بشكل يتناسب مع حجم الطبقة في المجتمع الإحصائي. ويعرف الأجراء التالي (بالمعاينة الطبقية التناسبية) التي يتم تطبيقها عندما تكون خواص المجتمع الإحصائي برمته هي النقطة الرئيسة في الدراسة. ويتم تمثيل الطبقة في العينة بنسبة مضبوطة، من تكرارها، في المجتمع الإحصائي كله. فمثلاً إذا كان 10% من المجتمع الذي قام بالتصويت هو من طلبة الكلية فإن 10% من عينة واحدة من الأصوات المراد استفتاءها سوف تؤخذ من هذه الطبقة.

(*) منته (Finite): مصطلح رياضي يشير إلى وجود عدد طبيعي محدد "ن" يدل على حجم المجتمع. (المراجع)

ومع ذلك، ينصب الاهتمام الرئيسي في بعض الدراسات البحثية على الفروقات بين الطبقات المتنوعة. ففي مثل هذه الحالات يختار الباحث عينات متساوية الحجم من كل طبقة. فمثلاً إذا كان أحدهم يبحث في الفرق بين عادات الدراسة لدى طلبة خريجين وطلبة كلية فإنه يقوم بتضمين أعداد متساوية لكلا المجموعتين، ثم يدرس الفروقات التي قد تكون موجودة بينهما. ويتم اختيار الأجراء المستخدم حسب طبيعة المسألة البحثية. فإذا كان التأكيد ينصب على أنماط الفروقات بين الطبقات، فإن المرء ينتخب أعداد متساوية من الحالات لكل منها. أما إذا كانت خواص المجتمع الإحصائي برمتها، هي الهدف الأساسي، فإن المعاينة المتناسبة تكون مناسبة أكثر.

وإذا كانت المعاينة الطبقية قابلة للتطبيق، فإنها تعطينا عينة أكثر تمثيلاً من المعاينة العشوائية البسيطة. ففي المعاينة العشوائية البسيطة، قد تكون بعض الطبقات، بطريق الصدفة، أكثر أو أقل تمثيلاً في العينة. فمثلاً، في العينة العشوائية البسيطة لطلبة الثانوية قد يكون من الممكن نظرياً (وهو مع ذلك بعيد الاحتمال) الحصول على إناث فقط. ولا يمكن، على أية حال، أن يحدث ذلك إذا تم إدراج الإناث والذكور بشكل مفصل وتم انتخاب عينة عشوائية من كل مجموعة.

وتتجلى المزية الكبرى للمعاينة الطبقية في كونها تضمن تمثيل المجموعات المحددة في المجتمع الإحصائي.

المعاينة العنقودية Cluster Sampling

كما ذكرنا سابقا، فإنه من الصعب، إن لم يكن من المستحيل، إدراج جميع أفراد مجتمع إحصائي مستهدف وانتخاب عينة منه. فمجتمع طلبة الثانوية الأمريكية مثلاً، كبير جداً بحيث لا يتسنى للمرء تسجيل جميع أفراده بهدف سحب عينة. وإضافة إلى ذلك، فإنه سيكون مكلفاً جداً دراسة عينة مشتتة في كل أنحاء الولايات المتحدة. ففي هذه الحالة سيكون من المناسب دراسة الأشخاص في مجموعات كما هي بصورة طبيعية أو في عناقيد / تجمعات. فمثلاً، قد يختار الباحث عدداً من المدارس بشكل عشوائي من قائمة مدارس، ثم يضمّن جميع الطلبة في تلك المدارس في العينة . إن هذا النوع من المعاينة الاحتمالية يشار إليه بالمعاينة العنقودية، لأن الوحدة المختارة ليست فرداً، بل مجموعة أفراد، هم معاً بصورة طبيعية. ويكوّن هؤلاء الأشخاص عنقوداً طالما أنهم متماثلون فيما يخص الخصائص المتعلقة بمتغيرات الدراسة. ولإيضاح ذلك، دعنا نفترض استفتاءاً للرأي العام أجري في أطلنطا. فقد لا يكون للباحث مجالاً للوصول إلى قائمة جميع السكان البالغين، وهكذا قد يستحيل اختيار عينة عشوائية بسيطة. أما الطريقة العملية الأخرى، فسوف تنطوي على انتخاب عينة عشوائية من، لنقل، خمسين مجمعاً من خارطة المدينة ثم القيام باستفتاء على جميع البالغين الذين يعيشون فيها. فكل مجمع يمثل عنقوداً من الأشخاص متشابهين في بعض الخصائص بارتباطهم بالعيش بجوار بعضهم. والتطبيق المألوف للمعاينة العنقودية يكون باستخدام صفوف تامة كعناقيد.

من المهم أن تكون العناقيد المتضمنة في دراسة معينة، قد انتخبت بصورة عشوائية من مجتمع إحصائي عنقودي. والشرط الإجرائي الآخر، هو أنه حينما يتم انتخاب العنقود فإن جميع أفراده يجب أن يكونوا في العينة. أما خطأ المعاينة في العينة العنقودية، فهو أكبر بكثير مما في المعاينة العشوائية. كما أن من المهم أن نتذكر أنه إذا كان عدد العناقيد صغيراً، فإن احتمال خطأ المعاينة يكون كبيراً - حتى وإن كان عدد الأشخاص كبيراً.

المعاينة المنتظمة Systematic Sampling

لا يزال هناك شكل آخر من المعاينة الاحتمالية يدعى بـ "المعاينة المنتظمة". وينطوي هذا الإجراء على سحب عينة وذلك بأخذ كل حالة ترتيبها (K) من قائمة المجتمع الإحصائي.

يقرر المرء أولاً مقدار عدد الأفراد الذين يريدهم في العينة (n). ولأنه يعرف العدد الكلي من الأفراد في المجتمع الإحصائي (N) فإنه ببساطة يقوم بتقسيم (N) على (n) ويحدد فترة المعاينــة (k) كي يطبقها على القائمة. ويتم اختيار الفرد الأول بصورة عشوائية من أفراد أول فترة (k) في القائمة، ثم يتم انتخاب كل فرد ترتيبه من مضاعفات (k) ابتداء من الفرد الأول من المجتمع الإحصائي لتكوين العينة. فمثلاً دعنا نفترض أن عدد أفراد مجتمع إحصائي 500 وحجم العينة المرغوبة 50 ، فتكون: $k = 500/50 = 10$.

قد يبدأ المرء قرب القمة في القائمة بحيث إن الحالة الأولى تنتخب بصورة عشوائية من أول عشر حالات، ثم يتم انتخاب كل حالة عاشرة بعدئذ. افترض أن الاسم الثالث أو العدد الثالث في القائمة كان الأول الذي تم انتخابه. عندئذ يضيف المرء فترة معاينة (k) أو (10) إلى (3) - لذا فالشخص الثالث عشر يقع في العينة وكذلك الثالث والعشرين وهكذا - وتستمر إضافة فترة المعاينة المتواصلة حتى بلوغ نهاية القائمة.

وتختلف المعاينة المنتظمة عن معاينة العينة العشوائية في كون أن الاختيارات المختلفة ليست مستقلة. فمتى ما اختيرت الحالة الأولى، فإنه سيتم تحديد جميع الحالات التالية في العينة بصورة أوتوماتيكية.

وإذا كانت قائمة المجتمع الإحصائي الأصلية مرتبة عشوائياً، فإن المعاينة المنتظمة سوف تعطينا عينة يمكن اعتبارها إحصائياً كبديل معقول لعينة عشوائية. غير أنه، إذا لم تكن القائمة عشوائية فإن من الممكن أن يكون لكل عضو ترتيبه (k) من المجتمع الإحصائي سمة فريدة قد تؤثر على المتغير التابع للدراسة، وتنتج عينة متحيزة. فالمعاينة المنتظمة من قائمة أبجدية مثلاً، قد لا تعطي عينة ممثلة لمجموعات قومية متنوعة لأن بعض المجموعات القومية تميل إلى التجمع تحت أحرف معينة، وقد تحذفها فترة المعاينة تماماً أو لا تحتويها، على الأقل، بمدى مناسب.

ويجدر الانتباه إلى أن الأنماط المختلفة من المعاينة الاحتمالية التي نوقشت، هي ليست مقصورة على جماعة بصورة متبادلة. فقد يتم استخدام توليفات متنوعة. فمثلاً قد يستخدم المرء

معاينة عنقودية إن كان يدرس مجتمعاً إحصائياً كبيراً جداً وواسع الانتشار. وفي نفس الوقت قد يهتم المرء بتصنيف العينة ليجيب على أسئلة تخص طبقاتها المختلفة. ففي هذه الحالة قد يصنف المرء مجتمعاً إحصائياً حسب الأصناف المحددة مسبقاً، ثم ينتخب عشوائياً بعدئذ، مجموعة الأفراد من بين كل طبقة.

المعاينة اللا احتمالية Nonprobability sampling

إن تعداد عناصر المجتمع الإحصائي في العديد من المواقف البحثية، وهو شرط أساسي في المعاينة الاحتمالية، يكون صعبا إن لم يكن مستحيلاً. ففي هذه الحالات يستخدم الباحث المعاينة اللا احتمالية التي تنطوي على إجراءات غير عشوائية لانتخاب أفراد العينة. أما الأشكال الرئيسية للمعاينة اللا احتمالية، فهي المعاينة العرضية، والمعاينة الهادفة، والمعاينة بالحصة / الحصصية (quota).

المعاينة العرضية Accidental sampling

تنطوي المعاينة العرضية التي تعد أضعف إجراءات المعاينة، على استخدام الحالات المتيسرة للدراسة. فمقابلة الأشخاص الأوائل الذين يواجههم المرء في حرم جامعي، أو استخدام الطلبة في الصف كعينة، أو إجراء مقابلة لمتطوعين في بحث مسحي، هي أمثلة متنوعة لمعاينة عرضية. وليست هناك طريقة (سوى تكرار الدراسة باستخدام المعاينة الاحتمالية) لتقدير الخطأ الناجم عن إجراءات المعاينة العرضية. ففي حالة استخدم المرء، المعاينة العرضية، عليه أن يكون حذراً جداً في تفسير النتائج.

المعاينة الهادفة/ القصدية Purposive Sampling

في المعاينة الهادفة، أو كما يشار إليها كذلك بمعاينة التحكيم، يتم اختيار عناصر المعاينة التي يحكم عليها بكونها نموذجية أو ممثلة، من المجتمع الإحصائي. والمعاينة الهادفة غالباً ما تستخدم للتنبؤ بالانتخابات القومية. ففي كل ولاية/ دولة يتم اختيار عدد من المقاطعات الصغيرة التي كانت نتائجها في الانتخابات السابقة نموذجية بالنسبة إلى الولاية كلها. فجميع أصحاب الأصوات المؤهلين في هذه المقاطعات، تتم مقابلتهم وتستخدم النتائج للتكهن بأنماط التصويت في الولاية. وباستخدام إجراءات مماثلة في كل الولايات فإن مستطلعي الاستفتاءات يتكهنون بالنتائج القومية.

إن السؤال المهم في المعاينة الهادفة هو المدى الذي يمكن فيه الاعتماد على الحكم للوصول إلى العينة النموذجية. وليس هناك من سبب للافتراض أن الوحدات التي حكم عليها بكونها نموذجية للمجتمع الإحصائي، ستستمر في كونها كذلك لفترة من الزمن. ونتيجة لذلك، فإن نتائج الدراسة باستخدام المعاينة الهادفة قد تكون مضللة.

وبسبب الكلفة المنخفضة، والراحة في الاستخدام، فإن المعاينة الهادفة تعتبر مفيدة في مسوحات الاتجاهات والآراء. وينبغي أن يعي المرء حدود واستخدام الطريقة بحذر شديد.

المعاينة الحصصية Quota Sampling

تنطوي هذه المعاينة على انتخاب حالات نموذجية من طبقات متنوعة للمجتمع الإحصائي. وتستند الحصص على خصائص معروفة في المجتمع الإحصائي، التي يرغب المرء في تعميمها. وتسحب العناصر بحيث تكون العينة الناتجة تقريباً مصغراً للمجتمع الإحصائي فيما يخص الخصائص المنتخبة. فمثلاً، إذا أوضحت نتائج الإحصاء السكاني أن 25 في المائة من المجتمع الإحصائي لمنطقة حضرية، يعيشون في الضواحي، فإنه يجب أن تأتي 25 في المائة من العينة من الضواحي.

وفيما يلي خطوات المعاينة بالحصة/ الحصصية:

1- تحديد عدد المتغيرات المرتبطة بشدة بالمسألة قيد الدرس، والمستخدمة كقواعد للتصنيف. فالمتغيرات كالجنس والعمر والثقافة والطبقة الاجتماعية غالباً ما تستخدم هنا.
2- يحدد استخدام الإحصاء السكاني أو أية بيانات متيسرة، حجم كل شريحة من المجتمع الإحصائي.
3- أحسب الحصص لكل شريحة من المجتمع بما يتناسب مع حجم كل شريحة.
4- انتخب الحالات النموذجية من كل شريحة أو طبقة، من المجتمع، لملئ الحصص.

ويكمن الضعف الكبير للمعاينة الحصصية في الخطوة (4)، أي في انتخاب الأشخاص من كل طبقة. فالمرء لا يعرف ببساطة ما إذا كان الأفراد المختارون يمثلون الطبقة المعينة أم لا. إن انتخاب العناصر يمكن أن يستند إلى سهولة الاختيار والفرصة المناسبة. فإذا قام أحدهم بانتخاب 25 في المائة من العوائل في مدينة داخلية لمسح معين، فالأكثر احتمالاً هو، لجوء المرء إلى العوائل الأكثر جذباً، وليست تلك المنفسحة، أو إلى تلك التي تكون سهلة المنال، أو إلى تلك التي يكون أفرادها فيها خلال اليوم، وهكذا. ومثل هذه الإجراءات تنتج تحيزاً منتظماً في العينة بشكل أوتوماتيكي لأن بعض العناصر سوف لن يتم تمثيلها. وعلاوة على ذلك، ليس هناك أساس لحساب الخطأ القائم في المعاينة الحصصية. ورغم هذه العيوب، تستخدم المعاينة الحصصية في العديد من المشاريع البحثية التي قد لا تكون ممكنة بخلاف ذلك. ويشعر العديدون بأن سرعة جمع البيانات تفوق المساوئ. وعلاوة على ذلك، فإن سنوات الخبرة مع المعاينة الحصصية جعلت من الممكن تحديد بعض الأخطاء، واتخاذ الخطوات في تجنبها.

ويمكن أن نجد نقاشاً آخر لأساليب المعاينة في أعمال (Cochran, 1985) و (sudman, 1976) المدرجة في قائمة المراجع في نهاية هذا الفصل.

حجم العينة (الجزء 1) The Size of The Sample (Part I)

غالباً ما يميل الناس العاديون إلى انتقاد البحوث (خصوصاً البحوث التي لا يحبون نتائجها) بقولهم إن العينة كانت أبسط من أن تسوغ استنتاجات الباحثين. لكن، ما الحجم الذي ينبغي أن يكون

للعينة؟ فإذا تساوت الأشياء الأخرى، تعتبر العينة الكبيرة أفضل تمثيلاً للمجتمع الإحصائي من العينة الصغيرة. ومع ذلك، فإن الخاصة الأهم للعينة، تكون في قدرتها على التمثيل وليس في حجمها. فعينة عشوائية من (200) هي أفضل من عينة عشوائية من (100) إلا أن عينة عشوائية من (100) هي أفضل من عينة متحيزة من (2,500,000).

وعليه، فإن الحجم وحده لا يضمن الدقة. فقد تكون العينة كبيرة لكنها لا تزال تحتوي على تحيز. ويتضح الوضع الأخير في استفتاء عام 1936 الذي تنبأ بهزيمة الرئيس روزفلت. فرغم أن العينة اشتملت على قرابة 2.5 مليون شخص، إلا أنها لم تكن ممثلة لأصحاب الأصوات، وهكذا تم الوصول إلى استنتاج خاطئ. إن التحيز قد نجم عن انتخاب المستجيبين للاستفتاء من تسجيلات السيارات، ومن أدلة الهاتف، وقوائم الاشتراك في المجلة. فهؤلاء الأشخاص لن يكونوا ممثلين لمجموع المجتمع الإحصائي المصوّت عام 1936. كذلك، فلأن الاستفتاء كان قد اجري بالبريد فقد كانت النتائج متحيزة عن طريق فروقات بين أولئك الذين استجابوا والذين لم يستجيبوا. وعليه، يجب أن يدرك الباحث أن حجم العينة سوف لن يعوض عن أي تحيز قد يدخل من خلال أساليب المعاينة الخاطئة. ويجب أن يبقى التمثيل الهدف الأساسي في انتخاب العينة.

ولاحقاً، في هذا الفصل، سوف ندخل إجراء تحديد حجم العينة المناسب على أساس مدى حجم التأثير الذي يكون ذا معنى، وكذلك على الاعتبارات الإحصائية . إن هذه الإجراءات التي تعرف (بحسابات القوة)، هي أفضل طريقة لتحديد أحجام العينات المطلوبة.

مفهوم خطأ المعاينة The Concept of Sampling Error

عندما يتم إجراء استدلال من عينة إلى مجتمع إحصائي فإن هذا سينطوي على مقدار معين من الخطأ لأنه، حتى، في العينات التي تكون عشوائية، يمكن أن نتوقع تغيرها من واحدة إلى أخرى. فالوسط الحسابي لدرجة الذكاء لعينة عشوائية واحدة بالنسبة لطلبة الصف الرابع قد تختلف عن الوسط الحسابي لدرجة الذكاء لعينة عشوائية أخرى من طلبة الصف الرابع، من نفس المجتمع الإحصائي. إن مثل هذه الفروقات التي تدعى (أخطاء المعاينة) تأتي من حقيقة أن المرء لاحظ عينة فقط، وليس المجتمع الإحصائي برمته.

ويعرف خطأ المعاينة على أنه "الفرق بين معْلم (*) / بارامتر المجتمع الإحصائي وإحصاة العينة". فمثلاً إذا كان المرء يعرف المتوسط الحسابي للمجتمع الإحصائي برمته (ويشار إليه بالرمز μ) وكذلك المتوسط الحسابي لعينة عشوائية (ويشار إليه بالرمز $\overline{X}$) من ذلك المجتمع الإحصائي فإن الفرق بين هذين الاثنين ($\overline{X}-\mu$) يمثل خطأ المعاينة (الذي يرمز إليه e) وهكذا،

(*) معْلم/ معْلمة: بارامتر (parameter): مصطلح يشير إلى خاصة لمجتمع إحصائي كالوسط الحسابي (μ) أو الانحراف المعياري (σ) مثلاً، ويقابل الإحصاءة (statistic) في العينة. وأحد أهداف الاستدلال، هو الاستدلال على المعْلم من الإحصاءة. المراجع

$e = \overline{X} - \mu$. فمثلاً إذا عرفنا أن الوسط الحسابي لدرجة الذكاء لمجتمع إحصائي مكون من 10,000 طالب في الصف الرابع هو μ=100 وإن عينة عشوائية معينة حجمها 200 لها وسط حسابي هو $\overline{X}$= 99 فإن خطأ المعاينة هو $\overline{X} - \mu$=99 - 100 = -1 . ولأننا نعتمد عادة على إحصاءات (جمع إحصاءة) العينات لتقدير معلمات/ بارامترات المجتمع الإحصائي، فإن مفهوم مدى تغير العينات من المجتمعات الإحصائية، هو عنصر أساسي في الإحصاء الاستدلالي. وعلى أية حال، بدلاً من محاولة تحديد التفاوت بين إحصاءات العينة ومعْلم/بارامتر المجتمع الإحصائي (الذي لا يكون معروفاً، غالبا) فإن الإجراء في الإحصاء الاستدلالي يكون بتقدير التباين المتوقع في إحصاءات عدد من العينات العشوائية المسحوبة من نفس المجتمع. ولأن كل إحصاءة في إحصاءات العينات تعتبر تقديراً لنفس معلم/بارامتر المجتمع الإحصائي، فإن التغير بين إحصاءات العينات يجب أن يعزى إلى خطأ المعاينة.

الطبيعة المطاوعة للقانون لأخطاء المعاينة
The Lawful nature of Sampling Errors

إذا أخذنا بنظر الاعتبار أن العينات العشوائية المسحوبة من نفس المجتمع الإحصائي ستختلف واحدة عن الأخرى، فهل أن استخدام عينة لأجراء استدلالات حول مجتمع إحصائي سيكون أفضل من مجرد التخمين؟ نعم، هو كذلك، لأن أخطاء المعاينة تتصرف بطريقة مطواعة للقانون وقابلة للتنبؤ. وتستمد القوانين الخاصة بخطأ المعاينة من المنطق الاستنتاجي، وتم التأكد منها من خلال الخبرة.

ورغم إنه لا يمكننا أن نتنبأ بطبيعة ومدى الخطأ في عينة واحدة، إلا أنه يمكننا أن نتنبأ بطبيعة ومدى أخطاء المعاينة بشكل عام. دعنا نوضح ذلك بالإشارة إلى أخطاء المعاينة المرتبطة بالمتوسط الحسابي.

أخطاء المعاينة للوسط الحسابي Sampling Errors of the Mean

يمكن دائماً توقع بعض خطأ المعاينة عندما يستخدم وسط العينة $\overline{X}$ لتقرير وسط المجتمع الإحصائي μ. ورغم إن مثل هذا التقدير يستند من حيث الممارسة على وسط عينة واحدة، فلنفترض أننا سحبنا عدة عينات عشوائية من نفس المجتمع الإحصائي، وقمنا بحساب الوسط لكل عينة. ستجد أن هذه الأوساط تختلف من واحد لآخر كما تختلف عن وسط المجتمع الإحصائي (إن كان معروفاً). إن هذا التغير بين الأوساط يعزى إلى خطأ المعاينة المرتبط مع كل عينة عشوائية كتقدير لوسط المجتمع الإحصائي. وقد جرت دراسة أخطاء المعاينة للوسط الحسابي بدقة ووجد أنها تتبع القوانين المعروفة.

The Expected mean of Sampling Errors Is Zero الوسط المتوقع لأخطاء المعاينة هو صفر: إذا أخذنا بنظر الاعتبار عدداً غير محدود من العينات العشوائية المسحوبة من مجتمع

إحصائي واحد، فإن الأخطاء الموجبة يتوقع لها أن توازن الأخطاء السالبة بحيث أن وسط أخطاء المعاينة سيكون صفراً. فمثلاً، إذا كان الوسط لأطوال مجتمع إحصائي من طلبة الصف الأول في الكلية هو 170 سم، وتم سحب عدة عينات عشوائية من ذلك المجتمع، فإننا نتوقع أن يكون لبعض العينات وسط أطوال، أكبر من 170 سم، بينما يكون للبعض الآخر وسط أطوال، أقل من 170 سم. وفي المدى البعيد، على أية حال، ستتوازن أخطاء المعاينة الموجبة والسالبة. فإذا كان لدينا عدد غير محدود من العينات العشوائية بنفس الحجم، وقمنا بحساب الوسط الحسابي لكل من هذه العينات ثم قمنا بحساب الوسط الحسابي لجميع هذه الأوساط، فإن وسط الأوساط سيكون مساوياً لوسط المجتمع الإحصائي.

خطأ المعاينة، تابع عكسي لحجم العينة Sampling Errors an inverse Function of Sample Size: كلما زاد حجم العينة، كان هناك تقلب أقل من عينة لأخرى في قيمة الوسط. وبعبارة أخرى، عندما يزداد حجم العينة يقل خطأ المعاينة المتوقع. والعينات الصغيرة هي أكثر عرضة لخطأ المعاينة من العينات الكبيرة. فقد يتوقع المرء أن الأوساط الحسابية القائمة على عينات من (10) تتقلب بصورة أكبر من الأوساط القائمة على العينات من (100). وفي مثالنا حول الأطوال، الأكثر احتمالاً، أن تشمل عينة عشوائية من (4 طلبة)، ثلاثة منهم فوق المعدل وواحد دون المعدل، مما هو في عينة عشوائية من (40 طالباً) فيها ثلاثون فوق المعـــدل و (10) دونه. فكلما ازداد حجم العينة، ازداد احتمال وجود وسط العينة قرب وسط المجتمع الإحصائي. فهناك علاقة رياضية بين حجم العينة وخطأ المعاينة. وسنبين لاحقاً كيف أن هذه العلاقة قد أدمجت في قوانين استدلالية.

خطأ المعاينة، تابع مباشر للانحراف المعياري للمجتمع Sampling Error Is a Direct function of the Standard Deviation of the Population: كلما كان لدينا انتشار أو تباين أكبر بين أفراد المجتمع الإحصائي، يكون توقعنا لانتشار أكبر في أوساط العينة. فمثلاً، إن وسط أوزان عينات عشوائية من (25) انتخبت كل منها من مجتمع المحترفين في سباق الخيول يبين مقدار خطأ معاينة أقل نسبياً من وسط أوزان عينات من (25) منتخبة من الفرسان من مجتمع إحصائي من المدرسين. فأوزان فرسان السباق المحترفين تقع داخل مدى ضيق بينما لا تقع أوزان المدرسين في نفس المدى. وعليه فبالنسبة إلى حجم عينة معينة، يكون خطأ المعاينة المتوقع لأوزان المدرسين، أكبر من خطأ المعاينة المتوقع بالنسبة لأوزان الفرسان.

أخطاء المعاينة موزعة بطريقة اعتيادية أو شبه اعتيادية حول الوسط الصفري المتوقع Sampling Errors Are Distributed in a Normal or Near-Normal Manner around the Expected Mean of Zero: ستكون أوساط العينات قرب وسط المجتمع الإحصائي بشكل متكرر أكثر من أوساط العينات البعيدة عن وسط المجتمع الإحصائي. وكلما

تحركنا أكثر فأكثر عن وسط المجتمع الإحصائي، نجد أوساط العينات أقل فأقل. وقد بينت النظرية والتجربة أن أوساط العينات العشوائية تتوزع بطريقة اعتيادية أو قريبة من الاعتيادية حول وسط المجتمع الإحصائي.

وبما أن خطأ المعاينة في هذه الحالة، هو الفرق بين وسط العينة ووسط المجتمع الإحصائي، فإن توزيع أخطاء المعاينة يكون في شكل اعتيادي أو قريب من الاعتيادي. ويكون التوزيعان بالتعريف، متطابقين سوى أن توزيع أوساط العينة، له وسط مساو لوسط المجتمع الإحصائي بينما وسط أخطاء المعاينة هو صفر.

سيشبه توزيع أوساط المعاينة المنحنى الاعتيادي، حتى وإن كان المجتمع الإحصائي الذي سحبت منه العينات غير موزع بصورة اعتيادية. فمثلاً، في مدرسة ابتدائية نموذجية نجد أعداداً متساوية من الأطفال حسب الأعمار المختلفة، ولذا فإن المضلع (Polygon) لأعمار الأطفال سيكون مستطيلاً بصورة أساسية. وإذا أخذنا عينات عشوائية من (40) تلميذاً من مدرسة ذات عدد متساو من الأطفال بعمر (6-11) لوجدنا العديد من العينات ذات وسط عمري قريباً من وسط المجتمع البالغ (8.5)، وأواسط العينات التي حجومها (8) أو (9) ستكون أقل شيوعاً بينما أوساط العينات التي هي إما متدنية (7) أو عالية (10) ستكون نادرة.

الخطأ المعياري للوسط الحسابي Standard Error of the Mean

طالما يمكن التنبؤ بالمدى والتوزيع لأخطاء المعاينة، فبوسعنا استخدام أوساط العينات بثقة تنبؤية لإجراء استدلالات تخص أوساط المجتمع الإحصائي. ومع ذلك، فإنا نحتاج إلى تقدير حجم خطأ المعاينة المرتبط بوسط العينة عندما يستخدم كتقدير لوسط المجتمع الإحصائي. والوسيلة المهمة لهذه الغاية، هي الخطأ المعياري للوسط الحسابي.

لقد تم إيضاح أن خطأ المعاينة يكشف عن نفسه في تغير أوساط العينات. وهكذا، إذا قام أحدهم بحساب الانحراف المعياري لمجموعة من الأوساط لعينات عشوائية من مجتمع إحصائي واحد، فسيكون لديه تقدير لمقدار خطأ المعاينة. ومن المحتمل، مع ذلك، الحصول على هذا التقدير على أساس عينة واحدة فقط. فقد رأينا أن هناك شيئين يؤثران على حجم خطأ المعاينة وهما: حجم العينة والانحراف المعياري في المجتمع الإحصائي. وعندما يعرف هذان الشيئان يمكن للمرء أن يتنبأ بالانحراف المعياري لأخطاء المعاينة. ويدعى هذا الانحراف المعياري المتوقع لأخطاء معاينة الوسط (بالخطأ المعياري للوسط) ويمثل بالرمز ($\sigma_{\bar{x}}$). لقد تم التوضيح من خلال المنطق الاستنتاجي أن الانحراف المعياري للوسط مساو للانحراف المعياري للمجتمع الإحصائي (σ) مقسوماً على الجذر التربيعي للعدد في كل عينة ($\sqrt{n}$). وبصيغة المعادلة:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad (6.1)$$

حيث:

$\sigma_{\bar{x}}$ = الخطأ المعياري للوسط.

σ = الانحراف المعياري للمجتمع الإحصائي.

n = العدد في كل عينة.

وفي الفصل الخامس، رأينا أن الانحراف المعياري (σ) هو مؤشر قياس لدرجة الانتشار بين الأفراد في مجتمع إحصائي. وبنفس الطريقة يكون الخطأ المعياري للوسط الحسابي ($\sigma_{\bar{x}}$) مؤشراً للانتشار المتوقع بين أوساط العينات المسحوبة بصورة عشوائية من مجتمع إحصائي. وكما سنرى فإن تفسير (σ) أو ($\sigma_{\bar{x}}$) متشابه.

ونظراً، لأن لأوساط العينات العشوائية توزيعات اعتيادية تقريبياً، فبوسعنا كذلك أن نستخدم نموذج المنحى- الاعتيادي لإجراء استدلالات تخص أوساط المجتمع الإحصائي. فإذا أخذنا بنظر الاعتبار أن الوسط المتوقع لأوساط العينات مساو لوسط المجتمع الإحصائي، وأن الانحراف المعياري لهذه الأوساط مساو للخطأ المعياري للوسط، وأن أوساط العينات العشوائية موزعة بصورة اعتيادية، فإنه يكون بإمكان المرء حساب الدرجة (z) لوسط عينة وإحالة (z) تلك، إلى جدول المنحنى الاعتيادي، لتقريب احتمال حدوث وسط العينة، من خلال الصدفة، بما يكون بعيداً أو أبعد من وسط المجتمع الإحصائي. ويتم الحصول على (z) بطرح وسط المجتمع الإحصائي من وسط العينة ثم تقسيم ذلك الفرق على الخطأ المعياري للوسط:

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \qquad (6.2)$$

ولإيضاح ذلك، خذ موظفة القبول في كلية، التي تتساءل ما إذا كان المجتمع الإحصائي للمتقدمين هو فوق المتوسط في الاختبار الفرعي اللفظي لامتحانات مجلس الكلية . إن الوسط القومي للدرجات اللفظية لمجلس الكلية هو (500) وإن الانحراف المعياري هو (100)، وهي تسحب عينة عشوائية من (64) من المجتمع الإحصائي وتجد أن وسط العينة هو (530). فتطرح السؤال: ما مدى احتمال سحب عينة عشوائية من (64) وذات وسط هو (530) من مجتمع إحصائي بوسط مقداره (500)؟ باستخدام المعادلة (6.1) تحسب موظفة القبول الخطأ المعياري للوسط بأنه 12.5:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{100}{\sqrt{64}} = 12.5$$

وبحساب درجة z لوسط عينتها، بالمعادلة (6.2) تحصل على النتيجة التالية:

$$z = \frac{\overline{X} - \mu}{\sigma_{\bar{x}}}$$
$$= \frac{530 - 500}{12.5}$$
$$= 2.4$$

وهكذا ينحرف وسط العينة عن وسط المجتمع الإحصائي ب (2.4) وحدات خطأ معياري، فما هو احتمال وجود وسط عينة ينحرف بهذا المقدار ($\sigma_{\bar{x}}$ 2.4) أو أكثر من وسط المجتمع الإحصائي؟ من الضروري فقط الرجوع إلى المنحنى الاعتيادي للتعبير عن هذا الانحراف (z) بصيغة الاحتمال. وبالرجوع إلى جدول المنحنى الاعتيادي، يجد المرء إن احتمال $z= 2.4$ أو ربما أعلى هو (0.0082). هذا يعني أن الدرجة z التي هي كبيرة أو أكبر، تحدث بالصدفة فقط ثمان مرات في كل (1000) مرة. ولأن احتمال وجود وسط العينة، من وسط المجتمع الإحصائي، على هذا البعد، فإن الموظفة تستنج أن وسط العينة، ربما لم يأت من مجتمع إحصائي وسطه الحسابي (500)، وعليه فإن وسط مجتمعها الإحصائي، أي المتقدمين لكليتها، هو ربما أكبر من (500).

استراتيجية الإحصاء الاستدلالي
THE STRATEGY OF INFERENTIAL STATISTICS

الإحصاء الاستدلالي، هو علم صنع القرارات المعقولة بمعلومات محدودة. فنحن نستخدم ما نلاحظه في العينات، وما نعرف عن خطأ المعاينة، للوصول إلى القرارات العرضة للخطأ، ولكنها معقولة، بالنسبة للمجتمعات الإحصائية. إن الإجراءات الإحصائية التي نقوم بها قبل هذه القرارات تدعى (اختبارات الدلالة). أما الأداة الأساسية لهذه الاختبارات الإحصائية فهي الفرضية الصفرية.

الفرضية الصفرية Null Hypothesis

أفترض أن لدينا (100) طالب من الصف الرابع للمشاركة في تجربة تخص تدريس بعض المفاهيم. ثم افترض أن فرضيتنا البحثية هي أن الطريقة (ب) للتدريس تؤدي إلى إتقان اكبر لهذه المفاهيم من الطريقة (أ). فنختار عشوائياً (50) طالباً ليتم تدريسهم هذه المفاهيم بالطريقة (أ) بينما الـ (50) طالب الآخر بالطريقة (ب). ونقوم بترتيب بيئتهم بطريقة تجعل المجموعتين تختلفان فقط في طريقة التدريس. ونجري في نهاية التجربة اختباراً يعتبر تعريفا إجرائيا مناسباً لإتقان عدد

من المفاهيم المعنية. ثم نجد أن الوسط الحسابي للطلبة الذين درسوا بالطريقة (ب) أعلى من الوسط لمن درسوا بالطريقة (أ). كيف نفسر هذا الفرق؟

وعلى افتراض أننا كنا دقيقين في جعل ظروف التعلم لكلا المجموعتين متكافئة، باستثناء طريقة التدريس، فبوسعنا أن نعزو الفرق بإعلان أن: (1) طريقة التدريس سببت الفرق أو (2) حدث الفرق بالصدفة. ومع أنه تم تخصيص الأفراد، بشكل عشوائي، لهذه المعالجات فإنه من الممكن، من خلال الصدفة، أن تكون تلاميذ مجموعة الطريقة (ب) أكثر ذكاءً وأعلى حافزية، أو ربما لسبب آخر كانوا أكثر احتمالاً لتعلم المفاهيم من الطلبة في مجموعة الطريقة (أ)، دون أهمية للطريقة التي تعلموا بها.

وعليه، فالفرق بين المجموعتين قد يكون نتيجة علاقة بين متغيرات – طريقة التدريس، وإتقان المفاهيم – أو ربما يكون نتيجة الصدفة وحدها (خطأ المعاينة). فكيف لنا أن نعرف أي إيضاح هو الصحيح؟ فلا يمكن، بالمعنى المطلق، أن نبرهن بصورة إيجابية أن طريقة التدريس قد سببت الفرق. وعلى أية حال بوسعنا أن نقدر احتمال الصدفة وحده بكونها المسؤولة عن الفرق الملاحظ ثم نحدد أي تفسير نقبله نتيجة لهذا التقدير.

ويعرف تفسير الصدفة (بالفرضية الصفرية)، التي كما تتذكر من الفصل الرابع، هي عبارة تقول بأنه لا توجد علاقة فعلية بين المتغيرات وإن أية علاقة ملحوظة، هي فقط من صنع الصدفة. وفي مثالنا توضح الفرضية الصفرية أنه لا توجد علاقة بين طريقة التدريس وإتقان عدد من المفاهيم.

وثمة طريقة أخرى لصياغة الفرضية الصفرية في مثالنا تنطوي على القول بأن الوسط الحسابي لجميع طلبة الصف الرابع الذين درسوا بالطريقة (أ) مساوٍ للوسط الحسابي لجميع طلبة الصف الرابع الذين درسوا بالطريقة (ب). وبصيغة المعادلة واستخدام الرمز (μ) للوسط الحسابي للمجتمع الإحصائي تصبح هذه العبارة:

$$H_o : \mu_A = \mu_B$$

حيث:

H_o = الفرضية الصفرية.

μ_A = الوسط الحسابي للطلبة الصف الرابع الذين درسوا بالطريقة (أ).

μ_B = الوسط الحسابي لطلبة الصف الرابع الذين درسوا بالطريقة (ب).

لاحظ أن الافتراض قد جرى ليفيد أن الـ (50) طالباً ممن درسوا حسب الطريقة (أ) هم عينة المجتمع الإحصائي لطلبة الصف الرابع الذين ربما درسوا حسب الطريقة (أ) وأن الـ (50) طالباً ممن درسوا حسب الطريقة (ب) هم عينة المجتمع الإحصائي الذين ربما درسوا حسب الطريقة

(ب). ويأمل الباحث استخدام البيانات من التجربة كي يستنتج ما يمكن توقعه عندما يتم تدريس طلبة الصف الرابع الآخرين بالطريقة (أ) أو (ب).

وفي تفسير الفرق الملاحظ بين المجموعتين، يجب أن يختار الباحث بين تفسير الصدفة (الفرضية الصفرية) والتفسير الذي مفاده أن هناك علاقة بين المتغيرات (فرضية البحث) - ويجب أن يقوم بذلك دون معرفة الحقيقة المطلقة التي تخص المجتمعات الإحصائية المعنية. ويستند الاختيار على معلومات غير كاملة، ولذا فإنه عرضة للخطأ المحتمل.

أخطاء النوع الأول (1) والنوع الثاني (2)
Type I and Type II Errors

إن الباحث سوف يستبقي أو يرفض الفرضية الصفرية. وأي قرار له سيكون إما صحيحاً أو خاطئاً. فإذا كانت الفرضية الصفرية صحيحة، فإن الباحث يكون على حق في الإبقاء عليها وعلى خطأ في رفضها. إن رفض فرضية صفرية صحيحة يدعى (خطأ النوع 1). وإن كانت الفرضية الصفرية خاطئة، يكون الباحث على خطأ في الإبقاء عليها، ومحقاً في رفضها. إن الإبقاء على فرضية صفرية خاطئة يدعى (خطأ النوع 2). وتتلخص الحالات المحتملة الأربعة في الجدول (6.1).

دعنا ندرس بعض النتائج المحتملة لنوعي الخطأ باستخدام مثالنا:

الجدول (6.1): تمثيل توضحي لأخطاء النوعين (1، 2)

		وضع حقيقي في المجتمع الإحصائي	
		H_0 صحيحة	H_0 خاطئة
قرار الباحث بعد	يرفض H_0	خطأ النوع 1	صحيح
إجراء اختبار الدلالة	يستبقى H_0	صحيح	خطأ النوع 2

خطأ النوع (1) Type I Error

يعتبر خطأ النوع 1 والذي يرمز إليه بالرمز ألفا الإغريقي (α) "تحذيراً كاذباً" - فالباحث يعتقد أن لديه شيئاً ما بينما لا يوجد هناك شيء. مثلاً، يرفض الباحث الفرضية الصفرية ويعلن أن هناك علاقة بين طريقة التدريس وإتقان المفاهيم العددية، وعليه يوصي بالطريقة (ب) على أنها أفضل. فتهمل المدارس الكتب والمواد القائمة على أساس الطريقة (أ) وتشتري مواد على أساس

الطريقة (ب). ويقام التدريب أثناء الخدمة لتدريب المدرسين على التدريس بالطريقة (ب). وبعد كل هذا الأنفاق في الوقت والمال، لا تلاحظ المدارس زيادة في إتقان المفاهيم العددية. فلا تعطي التجارب التالية، النتائج الملاحظة في الاستقصاء الأصلي. فرغم أن الحقيقة النهائية أو الخطأ للفرضية الصفرية يظلان مجهولين، فإن الدليل الداعم كبير. ولسوف يرتبك الباحث الأصلي ويشعر بالإهانة.

خطأ النوع (2) Type II Error

إن خطأ النوع (2) الذي يرمز له بالرمز الإغريقي بيتا (β) يُعد "خسارة" - فالباحث يستنتج أنه لا يوجد شيء، بينما هناك شيء ما حقاً. وفي مثالنا يستنتج الباحث أن الفرق بين المجموعتين قد يعزى إلى الصدفة، وأن الفرضية الصفرية على الأرجح صحيحة. وهكذا يحتفظ الباحث بالفرضية الصفرية، ويعلن أن هناك دليلاً غير كاف للاستنتاج بأن طريقة ما، هي أفضل من الأخرى. ويستنتج الباحثون الآخرون أن الطريقة (ب) هي أفضل من الطريقة (أ)، وأن المدارس التي تغير من الطريقة (أ) إلى الطريقة (ب) تسجل فوائد بارزة في إتقان الطلبة. ورغم أن الحقيقة المطلقة لا تزال مجهولة، إلا أن عدداً وافراً من الأدلة يدعم الفرضية البحثية. وسيشعر الباحث الأصلي بالإرباك (لكنه لا يشعر على الأرجح بالإهانة).

مقارنة أخطاء النوع (1) والنوع (2)

Comparison of Type I and Type II Errors

تؤدي أخطاء النوع (1) إلى تغيرات غير مسوغة، بينما تؤدي أخطاء النوع (2) ، نموذجياً، إلى الاحتفاظ بالوضع الراهن عندما تتوفر مسوغات للتغيير. فنتائج خطأ النوع (1) تعتبر عموماً أكثر خطورة من نتائج خطأ النوع (2) رغم وجود بعض الاستثناءات.

مستوى الدلالة LEVEL OF SIGNIFICANCE

تذكّر أن جميع الاستنتاجات العلمية هي عبارات ذات احتمال عال من الصحة، وليست عبارات مطلقة الصحة. ولكن ما مدى ارتفاع الاحتمال الذي يجعل الباحث مستعداً لإعلان وجود علاقة بين المتغيرات؟ وبعبارة أخرى، ما مدى ضآلة احتمال صحة الفرضية الصفرية قبل أن يرفضها الباحث؟ وتتنوع نتائج رفض فرضية صفرية صحيحة، خطأ النوع (1)، مع الوضع. وعليه يتمعّن الباحثون عادة في النتائج النسبية لأخطاء النوع (1) والنوع (2) ثم يقررون قبل إجراء تجاربهم حول مدى قوة الدليل قبل أن يرفضوا الفرضية الصفرية. إن هذا المستوى السابق للقرار الذي يتم فيه رفض الفرضية الصفرية، يدعى (مستوى الدلالة). فاحتمال خطأ النوع (1) يقع مباشرة تحت سيطرة الباحث الذي يقرر مستوى الدلالة حسب نوع الخطأ الذي يرغب في

تجنبه.

وبالطبع، يمكن للمرء أن يتجنب أخطاء النوع (1) بالاحتفاظ دوما بالفرضية الصفرية أو تجنب أخطاء النوع (2) برفضها دوماً. إلا أن أياً من هذه البدائل غير مثمر. فإذا كانت عواقب خطأ النوع (1) خطيرة جداً وعواقب خطأ النوع (2) ضئيلة، فقد يقرر الباحث المجازفة بإمكانية الخطأ (1) في حالة ما إذا كان يعزى الاحتمال المقدر لحدوث العلاقة الملاحظة، إلى مجرد الحظ، هو مرة بين ألف أو أقل. ويدعى هذا، اختبار للفرضية على مستوى دلالة (0.001) ويعتبر مستوى متحفظاً. وفي هذه الحالة يكون الباحث حريصاً في عدم الإعلان عن وجود علاقة حينما لا توجد علاقة البتة. ومع ذلك، فإن هذا القرار يعني قبول احتمال عال لخطأ النوع (2) بإعلان أنه لا توجد علاقة، بينما توجد علاقة في الواقع.

إذا كان التقدير بأن عواقب الخطأ (1) غير خطيرة، فقد يقرر الباحث الإعلان بأن العلاقة موجودة إذا كان احتمال وجود العلاقة الملحوظة بسبب الحظ فقط، هو مرة واحدة من عشر أو أقل. ويدعى ذلك اختبار الفرضية على مستوى دلالة (0.10). وهنا يتخذ الباحث احتياطات متواضعة فقط إزاء خطأ النوع (1) لكنه مع ذلك، لا يجازف كثيراً بخطأ النوع (2).

إن مستوى الدلالة هو احتمال الخطأ (1) الذي يجعل الباحث مستعداً للمجازفة برفض الفرضية الصفرية. وإذا قرر الباحث مستوى الدلالة عند (0.01) فهذا يعني أن الفرضية الصفرية سترفض إذا كان الاحتمال المقدر لوجود العلاقة الملحوظة، كحادث صدفة، هو واحد من مائة. وإذا تقرر مستوى الدلالة عند (0.0001) فإن الفرضية الصفرية سترفض في حالة ما إذا كان الاحتمال المقدر لوجود العلاقة الملحوظة بوصفها مجرد صدفة هو واحد من (10,000) أو أقل. إن مستويات الدلالة الأكثر شيوعاً في ميدان التربية هي مستويات (0.05) و (0.01).

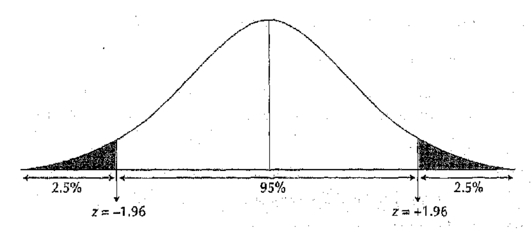
ويحدد الباحثون، بصورة تقليدية، مستوى الدلالة بعد تدقيقهم بالخطورة النسبية لأخطاء النوع (1) و (2)، وذلك قبل إجراء التجربة. فإذا كانت البيانات المستمدة من التجربة الكاملة توضح أن احتمال صحة الفرضية الصفرية مساو أو أقل من الاحتمال المقدر مسبقاً والمقبول، ترفض الفرضية الصفرية وتعلن النتائج في أنها دالة إحصائياً. وإن كان الاحتمال أكبر من الاحتمال المقدر مسبقاً والمقبول، فسوف توصف النتائج بأنها ليست ذات دلالة – أي أنه يتم الاحتفاظ بالفرضية الصفرية.

إن المعنى المألوف لكلمة (significeant) هو "مهم" أو "ذو معنى" واصطلح على ترجمتها بكلمة "دلالة" في مجال البحث العلمي. وفي الإحصاء تعني الكلمة "احتمال الصدفة أقل من احتمال مقدر سلفاً". إن نتائج البحوث يمكن أن تكون دالة إحصائياً دون أن تكون ذات معنى بصورة متأصلة أو مهمة.

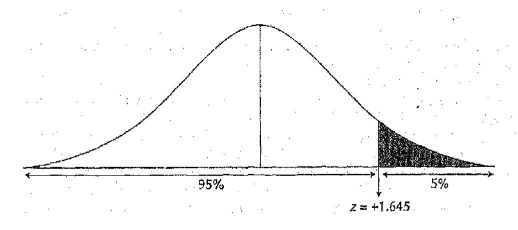
الاختبارات الموجهة وغير الموجهة Directional and Nondirectional Tests

في اختبار الفرضية الصفرية، لا نهتم عادة باتجاه الفروقات. بل نحن معنيون بمعرفة الابتعاد المحتمل لإحصاءات العينات عن معْلمات / بارامترات المجتمع الإحصائي. ففي المثال السابق لموظفة القبول، كان اهتمامنا الرئيس ينصب على تحديد ما إذا كان الفرق بين وسط العينة ووسط المجتمع الإحصائي قد تجاوز المقدار الذي يمكن أن يحدث بالصدفة. وكان هذا النوع من الاختبار يدعى بالاختبار (غير الموجه) لأن الباحث كان مهتماً بالفروقات في كل اتجاه. ويذكر الباحث فقط، أنه سيكون هناك فرق. لاحظ في الشكل (6.1) أن منطقة الرفض مقسمة بالتساوي بين طرفي التوزيع. وهكذا، إذا لوحظ أن الوسط الحسابي لعينة هو أما أكبر أو أصغر بشكل كاف من القيمة المفترضة، فإن الفرضية الصفرية سوف ترفض. إن اتجاه الفرق غير مهم.

الشكل (6.1): منحيات المناطق الحرجة للاختبارات الموجهة وغير الموجه



أ. منحنى بين المنطقة الحرجة لاختبار غير موجه (اختبار بطرفين)



ب. منحنى يبين المنطقة الحرجة لاختبار موجه (اختبار بطرف واحد)

وعلى أية حال، إن كان هناك بديل واحد فقط للفرضية الصفرية ذو دلالة، عندها سوف يستخدم الاختبار الموجه. مثلاً، إن الباحث الذي يدرس أثار حمية معينة على أناس بدينين، سيكون مهتماً فقط بتقدير احتمال أن الحمية تنقص الوزن. أو، إن كان مدرب كرة السلة يستقصي تأثير نظام تمرين على معدل النجاح في الرمي من الخط الخارجي، فإنه سوف ينفذ النظام فقط، إن كان لديه دليل معقول أنه يزيد النجاح. أما النتائج التي توضح أن هذا النظام يقلص النجاح، فإنها ستؤدي إلى نفس القرار تماماً كما تؤدي ذلك الفرضية الصفرية المستبقاة: إنس النظام .

إذا اختار الباحث، استناداً إلى التجربة أو البحوث السابق أو النظرية، الكشف عن اتجاه الفروقات المحتملة فإنه سيجري اختباراً موجهاً. وتوضح الفرضية الموجهة عما إذا كان المعْلم / البارامتر أكبر أو أقل من القيمة المفترضة. وهكذا، ففي الاختبارات الموجهة تقع المنطقة الحرجة، فقط، في واحد من طرفي التوزيع. فبالنسبة إلى مستوى دلالة (0.05) تكون هذه المنطقة في المنحنى الاعتيادي هي الدرجة المساوية لـ z = 1.645. ولا نقسم، كما نفعل في الاختبارات غير الموجهة، الـ (0.05) بين جانبي المنحنى. بل نضع كل الخمسة في المائة لخطأ الصدفة على جانب واحد من المنحنى. وهذا يعني أن الفرضية الصفرية سوف يتم الاحتفاظ بها ما لم يكن الفرق الملاحظ في "الاتجاه المفترض". ومن الواضح فإن رفض الفرضية الصفرية على مستوى دلالة معين يتطلب اختبار موجها تكون فيه قيمة z أصغر من اختبار غير موجه (قارن z = 1.645 مع z = 1.96). لذا، فإن الاختبار الموجه يجعل من السهل رفض الفرضية الصفرية ومن ثم يزيد احتمال رفضها إذا كان الفرق في الاتجاه المفترض.

وبالمصطلحات الإحصائية، يشار في الغالب إلى الاختبار غير الموجه يكونه اختبار دلالة بطرفين، بينما يشار للاختبار الموجه بطرف واحد. والقرار باستخدام أي منهما يجب أن يتم في وقت مبكر من الدراسة، وذلك قبل إجراء الاختبارات الإحصائية ولا ينتظر المرء أن يرى ماذا ستكون حال البيانات، ثم يختار اختباراً بطرف أو بطرفين.

حجم العينة (الجزء 2) The Size of The Sample (Part II)

الآن، وبعد أن درسنا أخطاء النوعين (1) و (2) ومستوى الدلالة، دعنا نَعُدْ إلى مسالة حجم العينة المطلوبة في الدراسة البحثية. إن الطريقة العلمية لتحديد حجم العينة المطلوب، يتم عن طريق تحديد حجم التأثير ذي المعنى (Δ)، ثم تحديد حجم العينة المطلوب للوصول إلى الاحتمال المرغوب لرفض الفرضية الصفرية على مستوى معين من الدلالة.

ولإيضاح ذلك دعونا نَعُدْ إلى موظفة القبول في الكلية التي تتساءل عما إذا كان مجتمعها الإحصائي من المتقدمين فوق المتوسط. ولتحديد عدد الأشخاص المطلوبين لاختبار فرضيتها، فإنها تحدد أولاً حجم تأثير قد يكون ذا معنى. فهي تقرر، مثلاً، إن حجم التأثير (0.33) (ثلث الانحراف المعياري) أو أكبر سيكون ذا معنى، إلا أن حجم تأثير أقل من (0.33) لن يكون ذا معنى.

إن تحديد ماهية حجم التأثير ذي المعنى هو قضية حكم مبرّر. وعلى أية حال، فإن المحترفين في حقولهم قادرون عادة على تحديد حجم التأثير الذي يفيد كخط فاصل معقول بين الفروقات ذات المعنى والتافهة، وفي هذا المثال، قد يعتبر حجم التأثير (0.33) من الاختبار اللفظي الفرعي للجنة القبول، خطأً فاصلاً معقولاً. إنه القول بأن درجة الوسط الحسابي (533) أو أكثر من [500 + (0,33) (100)] في اختبار مجلس القبول تكفي بأهميتها، بينما حجم التأثير الأصغر له أهمية قليلة أو ليس له أهمية عملية البتة. وبالرجوع إلى جدول المنحى الاعتيادي (الجدول A.1 في الملحق)، نرى أن للدرجة z (0.33) لها نسبة مئوية 62.93 تقع دونها. وعليه فإن موظفة القبول تقول أنه إذا كان لدى قرابة 63 في المائة من المجتمع الإحصائي العام درجات أقل من المتوسط لمجتمعها من المتقدمين، فإن الفرق يعتبر ذا معنى بينما الفرق الأقل لا يعد ذا معنى.

وبعدئذ تقرر موظفة القبول مستوى الدلالة لديها ويشار إليه بالرمز ألفا (α) واحتمالها المرغوب لرفض الفرضية الصفرية ($1-\beta$). دعنا نقل أنها تقرر بأنها تريد فرصة 90 في المائة لرفض الفرضية الصفرية مع طرف واحد وألفا (0,05).

ولديها الآن جميع المكونات المطلوبة لتحديد حجم العينة المطلوب، كما هو مبين في المعادلة التالية:

$$N=\left(\frac{1}{\Delta}\right)^2 (z\alpha + z\beta)^2 \qquad (6.3)$$

حيث:

N = العدد المطلوب في العينة.
Δ = حجم التأثير المحدد.
$Z\alpha$ = الدرجة z لمستوى الدلالة.
$Z\beta$ = الدرجة z للاحتمال المطلوب لرفض الفرضية الصفرية.

في الجدول A.1 نجد الدرجة z لألفا (0.05) ذي الطرف الواحد هي (1.654). وأن الدرجة z لاحتمال ذي طرف واحد لـ 90 في المائة (الاحتمال المطلوب لرفض الفرضية الصفرية) هو (1.28) وباستبدال قيم z و Δ المحددة ، نجد:

$$\begin{aligned} N&=\left(\frac{1}{.33}\right)^2 (1.645+1.28)^2 \\ &=3^2\,(2.925)^2 \\ &=9(8.56) \\ &=77.04 \end{aligned}$$

ومع عينة عشوائية من (78) شخصاً من مجموع مجتمعها الإحصائي للمتقدمين، يكون لموظفة القبول فرصة 90 في المائة لرفض الفرضية الصفرية على مستوى (0.05) بطرف واحد إن كان حجم التأثير الحقيقي في مجتمعها الإحصائي هو (0.33) أو أكثر.

لاحظ أنه كلما كان حجم التأثير صغيراً. كان العدد المطلوب أكبر. فإذا أرادت الباحثة حجم التأثير من (10) كي تحصل على احتمال 90 في المائة لرفض الفرضية الصفرية بطرف واحد ومستوى (0.05) فإنها سوف تحتاج إلى 856 فرداً في عينتها. لاحظ كذلك أنه كلما ازداد (1-β) وهو الاحتمال المطلوب لرفض الفرضية الصفرية، وكلما صغرت ألفا (مطالب كثيرة)، فإن عدد الأشخاص المطلوب يصبح أكبر.

إن تحديد العدد المطلوب في عينة يعتمد على مدى الدقة التي تريدها – أي مدى كبر أو صغر حجم التأثير الذي تريده أن يكون دالاً إحصائياً، ومقدار فرصة خطأ النوع (1) الذي ترغب في القبول به، ومقدار احتمال الرفض الذي تريده لفرضية صفرية خاطئة. وهذه كلها أحكام مسوغة، لكنه يمكن لجميعها أن تتم على أساس منطقي / معقول. وسندافع عن المواصفات، Δ= 0.33, α =.05 و 90.= (1-β). في المثال بكونها معقولة. وعليه نعتبر حجم عينة (78) كاف في هذه الحالة. ورغم أنه قد يكون بدعة اقتراح عينة أكبر ما يمكن، غير أنه من الصحيح أنه كلما زاد حجم العينة، فإن أحجام التأثير التي تتجه نحو الصغر، تصبح دالة إحصائيا. ونوصي أن يؤخذ بنظر الاعتبار، عند تخطيط البحوث، حجم التأثير والدلالة الإحصائية.

إن المعادلات، كالمعادلة (6.3) تدعى "معادلات القوة". وهي تشير إلى عدد الأشخاص المطلوب لإعطاء الباحث القوة المرغوبة لرفض الفرضية الصفرية إزاء حجم تأثير ومستوى دلالة محددين. فالعديد من كتب الإحصاء توضح حسابات تحديد الأعداد المطلوبة لدى تحديد حجم التأثير في إطار التناسب أو الارتباط أو الإحصاءات الأخرى. ويقدم (Borenstein & Cohen: (Statistical Power Analysis, 1989) معلومات موسعة حول هذه الإجراءات.

اختبار "ت" THE t-TEST

من بين أكثر الطرق استخداماً لاختبار الفرضيات الصفرية هي اختبارات "t". لقد أوضحنا أن من الممكن الاستفادة من منحنى الاحتمال الاعتيادي لمقارنة الوسط الحسابي لعينة مع الوسط الحسابي للمجتمع الإحصائي باستخدام الدرجة z لنرى ما إذا كان الوسط الحسابي للعينة ممثلاً للوسط الحسابي للمجتمع الإحصائي. وفي مثالنا، وجدت موظفة القبول الخطأ المعياري للوسط الحسابي لتوزيع العينة، ثم استخدمت المعادلة $(\overline{x}-\mu)/\sigma_{\overline{x}}$. وما يتضمنه استخدام هذا الإجراء هو ملاءمة منحنى الاحتمال الاعتيادي.

لقد تم الإيضاح رياضيا أن المنحنى الاعتيادي يكون مناسباً لاختبار الفرضية. فقط، عندما يعرف الانحراف المعياري للمجتمع الإحصائي كما هو الحال مع درجات مجلس الكلية. وفي أغلب الأوضاع البحثية لا يكون الانحراف المعياري للمجتمع الإحصائي معروفاً، ولكن يمكن تقديره عن طريق المعادلة التالية:

$$s=\sqrt{\frac{\sum x^2}{n-1}} \qquad (6.4)$$

حيث:

s = الانحراف المعياري التقديري للمجتمع الإحصائي.

Σx^2 = مجموع مربعات انحراف الدرجات، وهو $\sum (x-\bar{x})^2$

n = عدد أفراد العينة.

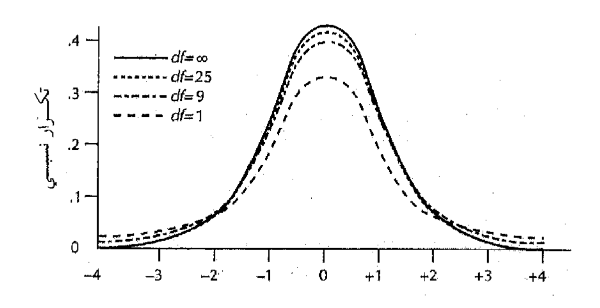
وعندما تستبدل هذه الـ (s) التقديرية، بالانحراف المعياري للمجتمع الإحصائي (σ) في حساب الخطأ المعياري للوسط الحسابي فإن من المعتاد التعبير عن المعادلة 6.1 كما يلي:

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}} \text{ بدلاً من } S_{\bar{x}}=\frac{s}{\sqrt{n}}$$

عندما تستخدم $S_{\bar{x}}$ بدلاً من $\sigma_{\bar{x}}$ فإن لكل حجم عينة منتهية توزيع احتمالي فريد. وتعرف هذه التوزيعات بمنحيات ت (t). وتصبح هذه التوزيعات مشابهة أكثر فأكثر للمنحنى الاعتيادي عندما يزداد حجم العينة. وقد تم تطوير سلسلة من التوزيعات تدعى توزيعات (t) لاختبار فرضيات تخص الوسط الحسابي للمجتمع الإحصائي باستخدام عينات صغيرة. وعندما يكون حجم العينة غير منته، يكون توزيع (t) مماثلاً للتوزيع الاعتيادي. وكلما أصبح حجم العينة أصغر، أصبح توزيع (t) مختلفاً بشكل متزايد عن توزيع (z). ولأجل أهدافنا، فإنه ليس من الضروري أن تعرف كيفية حساب توزيعات (t) لأن معظم النتائج المطلوبة ، على الأغلب، من هذه الحسابات يمكن إيجادها في الجدول A.2 في الملحق. إن منحنى (t) لا يقترب من خط الأساس بالسرعة التي يقترب بها المنحنى الاعتيادي. وبعض منحنيات (t) مبينة في الشكل (6.2) مع المنحنى الاعتيادي والخط الغامق المسمى $df=\infty$.

وتسمى منحنيات (t) حسب درجات حريالها وتختصر بالرمز (df). وقبل المضي في نقاش أكثر حول خواص منحنيات (t)، دعنا نتحول إلى مفهوم درجات الحرية.

الشكل (6.2): (منحنيات t لدرجات حرية مختلفة)



درجات الحرية Degrees of Freedom

يشير عدد درجات الحرية إلى عدد الملاحظات الحرة في تغيرها حول مُعلم / بارامتر ثابت. ولإيضاح المفهوم العام لدرجات الحرية، افترض أن مدرساً سأل طالبا أن يسمي له أية خمسة أعداد تخطر في باله. عندئذ سيكون الطالب حراً في تسمية أية خمسة أعداد يختارها وهذا سنقول إن للطالب خمس درجات حرية. افترض الآن إن المدرس سأل الطالب أن يسمي خمسة أعداد شريطة أن يكون الوسط الحسابي لهذه الأعداد الخمسة (20). فالطالب، الآن، حر في تسمية أية أعداد لأول أربعة، لكن تسمية العدد الأخير تتطلب أن يكون مجموع الخمسة أعداد (100)، بغية الوصول إلى الوسط الحسابي (20). فلو قام الطالب بتسمية الأعداد الأربعة الأولى (35، 20، 16، 10) فلابد أن يكون الرقم الخامس هو (19). فللطالب خمسة أعداد لتسميتها مع قيْد واحد، لذا فإن درجات حريته هي 4 = 1-5. ويمكن أن نبين هذه المعادلة:

$$
\begin{aligned}
df &= n - 1 \\
&= 5 - 1 \\
&= 4
\end{aligned}
$$

أفترض الآن، أن المدرس سأل الطالب أن يسمي سبعة أعداد بحيث يكون للثلاثة الأولى منها

وسط حسابي هو (10) وإن لجميع السبعة أعداد وسط حسابي هو (12). هنا لدينا سبعة أعداد وقيدان، لذا.

$$df = n - 2$$
$$= 7 - 2$$
$$= 5$$

عندما يتم تقدير الانحراف المعياري للمجتمع الإحصائي المجهول من الانحراف المعياري للعينة $\left(s_{\bar{x}} = \frac{s}{\sqrt{N}} \right)$ فستفقد درجة حرية واحدة، فدرجة حرية واحد من (df) ستفقد لأن إحصاءة العينة (s) تستمد من الانحرافات حول وسط العينة التي يكون مجموعها دوماً مساوياً للصفر. وهكذا فإن جميع الانحرافات عدا واحد، حرة في التغير أو (df = N-1).

إن مفهوم درجات الحرية موجود في أغلب الإجراءات في الإحصاء الاستدلالي. وهناك طريقة مناسبة لحساب درجات الحرية المرتبطة بكل إجراء. فبشكل عام، يعتمد عدد درجات الحرية، الذي تستند إليه إحصاءة العينة، على حجم العينة (N) وعلى عدد إحصاءات العينة المستخدمة في حسابه.

اختبار ت للعينات المستقلة

The t-Test for Independent Samples

غالباً ما يستخدم الباحثون عينتين عشوائيتين من مجتمع إحصائي ويخصصون معالجة مختلفة لكل مجموعة. وبعد التعرض لهذه المعالجات، تقارن المجموعتان بالنسبة للمتغير / المتغيرات التابعة لتقدير أثر المعالجات. والفرق الملاحظ بين المجموعتين، بعد مثل هذه المعالجات، قد لا يكون دالاً إحصائياً – أي يعزى إلى الصدفة – وعلى الباحثين أن يحددوا ما إذا كانت تلك هي الحالة. أما المؤشر المستخدم لإيجاد دلالة الفرق بين أوساط العينتين لهذا الهدف، يدعى اختبار (t) للعينات المستقلة. ويشار إلى هذه العينات بأنها مستقلة، لأنه جرى اختيارها بشكل مستقل من مجتمع إحصائي دون مزاوجة أو أية علاقة بين المجموعتين.

دعنا نستخدم مثالاً على ذلك. افترض أن باحثاً كان مهتماً بإيجاد ما إذا كان التوتر يؤثر على أداء حل المشكلة. الخطوة الأولى هي تعيين طلبة مقرر دراسي، عشوائياً، في مجموعتين، ولأنه جرى تحديد أفراد المجموعتين بشكل عشوائي فإن الوسطين الحسابيين لأداء المجموعتين في مهمة حل مشكلة لا ينبغي أن يختلف بشكل دال قبيل المعالجة. وبعد المعالجة، على أية حال، ينبغي أن يختلف الوسطان الحسابيان لأداء كلتا المجموعتين بشكل دال إذا ارتبط التوتر حقاً بأداء حل المشكلة.

ويبين الجدول 6.2 درجات حل المشكلات بعد المعالجة (X) ودرجات الانحراف (x) ومربعات درجات الانحراف (x^2) لأفراد المجموعتين التي تعمل إحداها تحت ظروف التوتر

والأخرى تحت ظروف الاسترخاء (لا توتر). إن درجة الوسط الحسابي لأداء الأفراد في مجموعة التوتر هو (10) ودرجة الوسط الحسابي لأداء مجموعة الاسترخاء هو (14). من الواضح أن هناك فرقاً. ونحتاج الآن إلى أن نحدد ما إذا كان الفرق يمكن أن يحدث بسهولة، بطريق الصدفة. ولأجل القيام بذلك يجب أن نقدر الفرق المتوقع بين المجموعتين من خلال الصدفة وحدها تحت فرضية صفرية حقيقية. والإجراء المناسب للقيام بذلك هو حساب الخطأ المعياري للفرق بين وسطين حسابيين ($S_{\bar{x}_1-\bar{x}_2}$)، والمعادلة لذلك في حالة العينات المستقلة هي:

$$S_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{\sum x_1^2+\sum x_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \qquad (6.5)$$

حيث:

$S_{\bar{x}_1-\bar{x}_2}$ = الخطأ المعياري للفرق بين الوسطين الحسابيين

n_1 = عدد الحالات في المجموعة 1

n_2 = عدد الحالات في المجموعة 2

$\sum x_1^2$ = مجموع مربعات درجات الانحراف في المجموعة 1

$\sum x_2^2$ = مجموع مربعات درجات الانحراف في المجموعة 2

إن الخطأ المعياري للفرق بين الوسطين يشار إليه أحياناً على أنه حد الخطأ لاختبار t، وفي مثالنا سيحسب ذلك كما يلي:

$$S_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{106+44}{15+15-2}\left(\frac{1}{15}+\frac{1}{15}\right)}$$

$$=\sqrt{\frac{150}{28}\left(\frac{2}{15}\right)}$$

$$=\sqrt{0.714}$$

$$=0.84$$

الجدول (6.2): حساب قيمة (t) لوسطي عينتين

مجموعة 1			مجموع 2		
X_1	x_1	x_1^2	X_2	x_2	x_2^2
18	+4	16	13	+3	9
17	+3	9	12	+2	4
16	+2	4	12	+2	4
16	+2	4	11	+1	1
16	+2	4	11	+1	1
15	+1	1	11	+1	1
15	+1	1	10	0	0
15	+1	1	10	0	0
14	0	0	10	0	0
14	0	0	10	0	0
14	0	0	10	0	0
13	-1	1	9	-1	1
12	-2	4	9	-1	1
11	-3	9	8	-2	4
10	-4	16	7	-3	9
8	-6	36	7	-3	9
$\frac{\Sigma X_1}{n_1}=\frac{210}{15}$		$\sum x_1^2=106$	$\frac{\Sigma X_2}{n_2}=\frac{150}{15}$		$\sum x_2^2=44$
$\overline{X}_1=14$			$\overline{X}_2=10$		

إن هذا الحساب يوضح لنا الفرق المتوقع من خلال الصدفة وحدها إذا كانت الفرضية الصفرية صحيحة. وبعبارة أخرى، إن القيمة (0.84) هي الفرق الذي نتوقعه بين الوسطين الحسابيين لدرجات الأداء لمجموعتينا إن كانتا مسحوبتين عشوائياً من مجتمع إحصائي عادي و لم تخضعا لمعالجات مختلفة. وإذا أخذنا بنظر الاعتبار عددا غير محدود من العينات في مثل هذه الظروف، فنتوقع أن نلاحظ الفرق بمعدل أقل من 0.84 (زائد أو ناقص) في 68 في المائة من حسابات الفروق بين مجموعات عشوائية وقيمة أكثر من 0.84 (زائد أو ناقص) في 32 في المائة الأخرى. (إنه خارج نطاق هذا النص، مناقشة سبب تطبيق المعادلة للخطأ المعياري للفرق بين الأوساط الحسابية بما ينتج عنه الفرق التقديري المناسب، العائد للصدفة) ولأجل مناقشة الأساس المنطقي لهذا الأجراء أنظر (Ary & Jacobs, 1976). وفي مثالنا لبيانات الجدول 6.2، ينبغي أن نتوقع

متوسط فرق مقداره (0.84) من خلال الصدفة تحت فرضية صفرية حقيقية. لقد لاحظنا فرقاً مقداره (4.0). فهل الفرق الملاحظ هو أكبر من الفرق المتوقع بما يكفي لرفض الفرضية الصفرية؟

للإجابة على هذا السؤال، نقوم أولاً بإعطاء نسبة للعددين. وهذه النسبة الذي تدعى النسبة (t) يعبر عنها بالمعادلة التالية:

$$t=\frac{\overline{X}_1-\overline{X}_2}{S_{\bar{x}_1-\bar{x}_2}} \qquad (6.6)$$

حيث:

t = النسبة t

$\overline{X}_1-\overline{X}_2$ = الفرق الملاحظ بين وسطين.

$S_{\bar{x}_1-\bar{x}_2}$ = الخطأ المعياري للفرق بين وسطين (الفرق المتوقع بين وسطين عندما تكون الفرضية الصفرية صحيحة).

يمكننا أن نكتب معادلة النسبة t بشكل أكثر اكتمالاً وذلك عن طريق تضمين معادلة الخطأ المعياري للفرق بين الوسطين.

$$t=\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\left(\frac{\sum x_1^2+\sum x_2^2}{n_1+n_2-2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \qquad (6.7)$$

في مثالنا عندئذ تكون قيمة النسبة t هي:

$$\frac{14-10}{0.84}=4.76$$

إن فرقنا الملاحظ هو (4.76) مرات لحجم الفرق المتوقع تحت فرضية صفرية صحيحة. فهل هو فرق كبير بما يكفي لرفض الفرضية الصفرية على مستوى (0.05)؟ وللإجابة على ذلك نحتاج فقط إلى حساب درجات الحرية والرجوع إلى جدول (t).

أما درجات الحرية لاختبار (t) المستقل، فهي عدد الحالات في كل مجموعة زائداً عدد الحالات في المجموعة الثانية، ناقصاً (2)، أو

$$df = n_1 + n_2 - 2$$

وفي مثالنا لدينا (15+15-2=28) درجة حرية. وبوسعنا الآن استخدام الجدول A-2 في الملحق

لتحديد دلالة نتائجنا. فالعمود الأول في هذا الجدول يرمز له df (درجات الحرية). ويجد المرء الصف المناسب في الجدول بتحديد درجات الحرية في دراسته. وبالنسبة لمثالنا نرجع إلى صف أل (28) درجة حرية. أما بقية الأعمدة فتبين قيم (t) المرتبطة باحتمالات معينة للاختبارات الموجهة وغير الموجهة. ولأن المتغير المستقل في مثالنا يمكن أن يؤثر على حل المشكلة باتجاه إيجابي أو سلبي فنحتاج إلى القيام باختبار غير موجه. وفي الصف لـ (28) درجة حرية نجد (1.701) في العمود الموسوم (0.10) للاختبار غير الموجه الذي يبين لنا أنه مع فرضية صفرية صحيحة و (28) درجة حرية لنسبة t تساوي (1.701+) أو أكثر أو (1.701-) أو أقل سيحدث بالصدفة مرة واحدة كل عشر مرات . إن العدد (2.048) في العمود الموسوم (0.05) يشير إلى فرضية صفرية صحيحة و (28) درجة حرية لنسبة (t) هي (2.048 ±) أو أكثر سيحدث بالصدفة في 5% من الوقت.

أما النسبة (t) التي لاحظناها وهي (4.76) أكبر من (2.048) بما يعني أن الفرق بين مجموعتينا أكبر من القيمة المطلوبة لرفض الفرضية الصفرية على مستوى الدلالة (0.05). إن الاحتمال المقدر للفرضية الصفرية الصحيحة هو أقل من 5 بالمائة ($P<.05$). ورغم أننا لا نعرف على وجه اليقين أن متغيرات التوتر وأداء حل المشكلات مترابطة، إلا أن الدليل دال بما فيه الكفاية حسب جملة أصنافنا السابقة بما يمكننا من الاستنتاج بأن العلاقة الملاحظة هي ربما ليست مجرد حادثة صدفة. فلو كانت النسبة (t) الملاحظة أقل من (2.048) لاستنتجنا أن الدليل لم يكن مناسباً بما يكفي كي يقودنا إلى بيان أن العلاقة موجودة بين المتغيرات. وبعبارة أخرى، نكون قد احتفظنا بالفرضية الصفرية.

لاحظ أننا كلما مضينا من اليسار إلى اليمين في الجدول (t) فإننا نجد قيم (t) المطلوبة لرفض الفرضية الصفرية على مستويات دقيقة بشكل متزايد للدلالة. وبالنسبة لـ (28) درجة حرية فإن قيمة (2.763) أو أكثر ستقود إلى رفض فرضية صفرية بمستوى (0.01). وقيمة (3.674) أو أكثر ستؤدي إلى رفض فرضية صفرية بمستوى (0.001). لذا، فإن القيمة (4.76) دالة، ليس فقط على مستوى (0.05) ($P<0.05$)، بل كذلك على مستوى (0.01) ($P<0.01$) ومستوى (0.001) ($P<0.01$).

ولأن المقام هو قياس مقدار الفرق الواضح الذي يمكن توقعه خلال الصدفة وحدها، فإنه يدعى حد الخطأ لاختبار (t). وإذا كانت النسبة للفرق الملاحظ (البسط) المقسوم على حد الخطأ (المقام) يساوي أو يتجاوز القيمة الموضحة في الجدول A.2 فبالوسع رفض الصفرية حسب المستوى الموضح من الدلالة.

وبتخصيص مجموعة اللاتوتر على أنها المجموعة الضابطة، يمكننا حساب حجم التأثير في مثالنا بقسمة الفرق بين الوسطين الحسابيين على الانحراف المعياري المقدر للمجتمع الإحصائي لمجموعة اللاتوتر:

$$S_c=\sqrt{\frac{\sum x^2}{n-1}}=\sqrt{\frac{106}{14-1}}=2.86$$

$$\Delta=\frac{\bar{X}_E-\bar{X}_C}{S_c}=\frac{14-10}{2.86}=1.40$$

إن الوسط الحسابي لحل المشكلة بالنسبة لمجموعة التوتر هو (1.4) انحراف معياري دون الوسط الحسابي لمجموعة اللاتوتر، والفرق جوهري ودال إحصائياً.

منطق اختبار ت The Logic of The t-Test

إن بسط اختبار (t) هو الفرق الفعلي الذي لوحظ بين المجموعتين، بينما المقام $S_{\bar{x}_1-\bar{x}_2}$ أو حد الخطأ هو تقدير مدى التوقع الذي تختلف فيه المجموعتان بالصدفة وحدها، أي أنه يوضح الفرق المتوقع بين مجموعتين منتخبتين عن طريق إجراء عشوائي من مجتمع إحصائي واحد. ويستند هذا المقام على 1) العدد في العينات، $(n_1 + n_2)$ (إذ كلما كان العدد أكبر، كانت الفروقات العشوائية المتوقعة بين أوساط العينات أقل) و 2) التباين داخل المجموعتين، S_2, S_1 (فكلما كان التباين داخل المجموعات أكبر، كانت الفروقات العشوائية المتوقعة بين المجموعات أكبر).

فإذا كانت نسبة الفرق الملاحظ (البسط) مقسوماً على حد الخطأ (المقام) يساوي أو يفوق القيمة الموضحة في جدول قيم "t" فإنه يمكن رفض الفرضية الصفرية حسب مستوى الدلالة الموضح.

الاختبار التائي (t) لعينات غير مستقلة
The t-Test for Nonindependent Samples

لقد تركز نقاشنا، حتى الآن، على مقارنة الوسطين اللذين حصلنا عليهما من عينتين مستقلتين ففي عينة مستقلة يتم اختيار كل عضو بشكل عشوائي من مجتمع إحصائي وتكوين مجموعة معينة ليس له صلة بتكوين المجموعة الأخرى. وعلى أية حال، قد يرغب الباحثون أحياناً أفراد مجموعتيهما حسب بعض الصفات المهمة لهدف البحث، أو قد يرغبون في مقارنة الوسطين الحسابيين للمجموعة ذاتها تحت ظرفين تجريبيين مختلفين. وفي مثل هذه الحالات لا تكون المجموعات مستقلة، بقدر ارتباط تكوين مجموعة معينة بتكوين المجموعة الأخرى. ونتوقع ترابطاً بين درجات المتغير التابع. لذا يجب استخدام الاختبار التائي "t"للأوساط غير المستقلة أو المترابطة. ويعرف هذا الاختبار أيضاً بالاختبار التائي المترابط أو التابع. أما القياس المراد تحليله بالاختبار التائي غير المستقل فهو الوسط الحسابي للفرق بين أزواج الدرجات.

دعنا ندرس أحد الأمثلة. أفترض أننا نرغب في معرفة ما إذا كانت دراسة مقرر في البحث العلمي ستؤثر على اتجاهات الطلبة نحو البحوث. ولاستقصاء ذلك، نختار صفاً يدرس مقرراً في

البحث العلمي ونحصل على قياسات اتجاهات الطلبة نحو البحث في أول وآخر يوم من الدوام. دعنا نفترض أننا جمعنا بعض البيانات وعرضت النتائج في الجدول (6.3). إن العمودين (2)، (3) يبينان درجات كل طالب في القياسيين الأول والثاني. ويعرض العمود (4) الفرق بين الدرجتين الأولى والثانية لكل طالب. ومجموع هذه الفروق تبلغ (30+). أما وسط الفروق (2+) فيتم الحصول عليه بتقسيم (30+) (ΣD) على N الذي هو عدد الملاحظات المزدوجة أو (15)، ويبين العمود (5) مربعات الفروق.

الجدول (6.3): درجات الاتجاه القبلي والبعدي
لـ 15 طالباً يدرسون مقرراً تمهيديا في البحث العلمي

(1) عدد الأفراد	(2) الاختبار القبلي	(3) الاختبار البعدي	(4) D	(5) D^2
1	10	12	+2	+4
2	9	13	+4	+16
3	8	12	+4	+16
4	11	9	-2	+4
5	10	8	-2	+4
6	7	9	+2	+4
7	10	12	+2	+4
8	9	11	+2	+4
9	8	10	+2	+4
10	6	10	+4	+16
11	10	12	+2	+4
12	7	13	+6	+36
13	10	6	-4	+16
14	9	13	+4	+16
15	10	14	+4	+16
			$\Sigma D = +30$	$\Sigma D^2 = +164$

أما معادلة الاختبار التائي غير المستقل فهي:

$$t = \frac{\overline{D}}{\sqrt{\dfrac{\Sigma D^2 - \dfrac{(\Sigma D)^2}{N}}{N(N-1)}}} \qquad (6.8)$$

حيث:

t = القيمة التائية للأوساط (المترابطة) غير المستقلة.

$\overline{D}$ = وسط الفروق.

D = الفرق بين كل أزواج الدرجات.

ΣD^2 = مجموع مربعات الفروق.

N = عدد الأزواج.

وباستبدال القيم من الجدول 6.3 نحصل على

$$t = \frac{\frac{30}{15}}{\sqrt{\frac{164 - \frac{(30)^2}{15}}{15(15-1)}}} = \frac{2}{\sqrt{\frac{164-60}{210}}} = \frac{2}{\sqrt{\frac{104}{210}}} = \frac{2}{\sqrt{0.4952}} = \frac{2}{0.704} = 2.84$$

وتدلنا النسبة التائية أن الفرق الملاحظ هو (2.84) مرة بحجم الفرق المتوقع حسب فرضية صفرية صحيحة. ويجب أن نعود الآن إلى جدول القيم التائية (الملحق A.2) لتحديد الدلالة الإحصائية للنسبة الملاحظة. إن عدد درجات الحرية للاختبار التائي غير المستقل هو N-1 حيث (N) تساوي عدد أزواج الملاحظات. وفي مثالنا نحصل على (15-1=14) درجة حرية. وفي جدول القيم التائية نجد أنه مع (14) درجة حرية هناك حاجة لقيمة تائية (2.145) وذلك كي يكون (t) دالاً على مستوى (0.05)، وقيمة تائية (2.977) للدلالة على مستوى (0.01) في حالة الاختبار غير الموجه. أما قيمة (2.84) فتتجاوز القيمة المعينة لمستوى (0.05) إلا أنها لا تصل إلى القيمة المعينة لمستوى (0.01). وهذا يعني أن الفرق بين الوسطين دال مستوى (0.05)، لكنه غير دال مستوى (0.01). ولو وضعنا مستوى الدلالة حسب (0.05) فيمكن أن نستنتج أن اتجاه الطلبة نحو البحث قد تغير.

الاختبار التائي لمعاملات ارتباط بيرسون (r)
The t-Test for Pearson r Correlation Coefficients

ثمة استخدام مهم آخر للاختبار التائي، هو في اختبار الفرضيات المتعلقة بالارتباط (ρ) وتُقرأ "رو" للمجتمع الإحصائي. أما الفرضية الصفرية الأكثر شيوعاً في مثل هذه الحالات فهي أن ارتباط الإحصائي هو (صفر) وإن الارتباط (r) الملاحظ في العينة هو من قبيل الصدفة. فمثلاً، قد يسحب الباحث عينة من (27) طالباً مستجداً في الكلية لفرض إجراء اختبارات المفردات والإدراك المكاني لهم، فوجد معامل بيرسون (r=0.25) بين القياسين. الخطوة التالية في تقرير ما إذا كان بالإمكان اعتبار معامل الارتباط الملاحظ نتيجة للصدفة في مجتمع إحصائي حيث الارتباط الحقيقي (ρ=0).

ولأجل اختبار الفرضية بأن معامل الارتباط هو (صفر) فليس عليك أن تحسب اختبار (t). فهذه الحسابات قد أجريت لأحجام عينات ومعاملات ارتباط متنوعة، وهي موضحة في الجدول A.5 في الملحق.

إن (r) الدالة هي التي تساوي أو أكبر من القيمة المجدولة مع N-2 درجة حرية، حيث N = عدد أزواج الدرجات. ويوضح الجدول A.5 أنه مع df = 25 (العدد في العينة ناقص 2) حيث يتم إجراء الاختبار ذي الطرفين، وعندها يتطلب أن يكون معامل بيرسون أعلى من (+.3809) أو أقل من (-.3801) لرفض الفرضية الصفرية على مستوى (0.05) ذي الطرفين. وهكذا فارتباط (0.20) الحاصل في الدراسة، الخاص بطلبة الكلية الجدد، ليس دالاً على مستوى الدلالة هذا. ومع عدد كبير من الحالات، قد يكون معامل الارتباط منخفضاً في القيمة إلا أنه قد يكون دالاً إحصائياً. ولأن قيمة الارتباط هي التي توضح درجة العلاقة بين المتغيرات، فإن ارتباطاً منخفضاً يشير دائماً إلى علاقة منخفضة حتى عندما يكون الارتباط دالاً إحصائياً. فمثلاً، وجد (Jackson & Lahadrine, 1967) مع عينة من (144) طالبة في الصف السادس ارتباطاً (+.25) بين استجابة الطالبات في استبيان لقياس رضى الطالبات بالمدرسة وتنبؤ مدرسيهن حول الكيفية التي تستجيب بها الطالبات للاستبيان. ولأن حجم العينة كبير (299) فإن الارتباط، رغم كونه منخفضاً، كان دالاً إحصائياً على مستوى (.01). وتشير هذه الاستنتاجات إلى أنه في المجتمع الإحصائي الممثل في العينة، ليس محتملا أن يكون الارتباط (0). وعلى أية حال، أن الارتباط المنخفض الملاحظ يوحي إلى أنه برغم أن بوسع المدرسين التنبؤ برضى الطالبات بما يفوق الصدفة إلا أن تنبوات المدرسين حول الرضى لها علاقة ضعيفة مع رضى الطالبات الفعلية.

ويمكن استخدام الاختبار "t" لاختبار الفرضيات حول معاملات ارتباط مجتمع إحصائي مختلفة عن (0). ويمكن كذلك استخدامه لاختبار الفرضية القائلة بأن الارتباطات الملاحظة في عينتين يمكن لها أن تبرز من نفس المجتمع الإحصائي. ونظراً لأن هذا الكتاب تمهيدي، فقد اخترنا عدم إدخال مثل هذه الاختبارات هنا. إن وصفاً مفيداً لهذه الاختبارات يمكن إيجاده في كتاب (Glass & Hopkins, 1984) وفي غيره من الكتب.

تحليل التباين ANALYSIS OF VARIANCE

في تحليل التباين (ANOVA)، كما في الاختبار، "t" تستخدم نسبة مقارنة الفروق الملاحظة مع حد الخطأ لاختبار الفرضيات حول الفروقات بين المجموعات. وهذه النسبة التي تدعى النسبة (F) تستخدم التباين (σ^2) لأوساط المجموعات كقياس للفروقات الملاحظة بين المجموعات. وهذا يعني أن "أنوفا" هو أسلوب أكثر تنوعاً من الاختبار التائي "t". فالاختبار التائي يمكن استخدامه فقط لاختبار الفرق بين وسطين. ويمكن لاختبار "ANOVA" أن يختبر الفرق بين وسطين أو

أكثر. ولا يستخدم بعض الإحصائيين الاختبار التائي مطلقاً، لأنه يمكن استخدام أنوفا في أي وضع يمكن فيه استخدام الاختبار التائي، ويمكن علاوة على ذلك القيام بعدة أشياء لا يستطيع الاختبار التائي القيام بها.

إن الأساس المنطقي العام لاختبار ANOVA يقوم على أن التباين الكلي لجميع الأفراد في تجربة معينة يمكن تقسيمه فرعياً إلى مصدرين: تباين بين مجموعات وتباين داخل مجموعات.

الجدول (6.4): القياسات المستحصلة في ثلاثة عينات عشوائية بعد أداء عمل تحت ظروف التوتر العالي والمعتدل وعدم التوتر

المجموعة 1 التوتر العالي		المجموعة 2 التوتر المعتدل		المجموعة 3 عدم التوتر	
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
19	361	22	484	15	225
18	324	20	408	14	196
17	289	19	361	14	196
16	256	18	324	13	196
15	225	17	289	13	169
15	225	16	256	12	144
14	196	16	256	12	144
13	169	15	225	11	121
12	144	14	196	11	121
11	121	12	144	10	100
$\Sigma X_1 = 150$	$\Sigma X_1^2 = 2310$	$\Sigma X_2 = 169$	$\Sigma X_2^2 = 2935$	$\Sigma X_3 = 125$	$\Sigma X_3^2 = 1585$
$\overline{X}_1 = 15.0$		$\overline{X}_2 = 16.9$	$\overline{X}_3 = 12.5$		$\overline{\overline{X}} = 14.8$
		$\Sigma X_t = 444$	$\Sigma X_t^2 = 6830$		

ويكون التباين بين المجموعات مشمولاً في البسط في النسبة (F). والتباين داخل المجموعات مشمول في حد الخطأ أو المقام، تماماً كما في الاختبار التائي. وكلما ازداد التباين بين المجموعات تزداد النسبة (F)، وإذا ازداد التباين داخل المجموعات، تناقصت النسبة F. ويؤثر عدد الأفراد على النسبة (F)، فكلما كان العدد كبيراً أصبح البسط كبيراً. وعندما يتساوى البسط والمقام لا تكون الفروقات بين أوساط المجموعة أكبر مما هو متوقع عن طريق الصدفة وحدها. وإذا كان البسط أكبر من المقام، يرجع المرء إلى جدول قيم (F) (مثل جدول A.3 في الملحق) لتحديد ما

إذا كان المعدل كبيراً بما يكفي لتمكين المرء من رفض الفرضية الصفرية على مستوى محدد مسبقاً.

حساب النسبة F (تحليل بسيط للتباين)
Computation of The F-Ratio (Simple Analysis of Variance)

افترض أن لدينا ثلاثة ظروف تجريبية للتوتر العالي والمعتدل وعدم التوتر وكنا نرغب في مقارنة أداء مجموعات ثلاث للأشخاص المعينين عشوائياً لهذه الظروف الثلاثة في واجب لحل مشكلة بسيطة. افترض أن البيانات المقدمة في الجدول 6.4 تلخص ملاحظاتنا لأداء هذه المجموعات الثلاث، وأننا بصدد اختيار الفرضية الصفرية بمعنى أنه ليس هناك فرق دال بين هذه الملاحظات على مستوى (01.).

ويمكن رؤية اختلاف الأوساط عن بعضها وعن الوسط الكلي لجميع الـ 30 فرداً ($\overline{\overline{X}}$ هو الوسط العام). فهل الفروقات بين هذه الأوساط كبيرة بما يكفي لتكون دالة إحصائياً أم هل إن من المحتمل أنها قد حدثت بالصدفة؟ للإجابة على تلك نحسب النسبة الفائية (F).

الخطوة الأولى هي إيجاد مجموع مربعات الانحرافات لكل من الدرجات الفردية عن الوسط العام. ويدعى هذا المؤشر بالمجموع الإجمالي للمربعات ويعكس جميع آثار المعالجة وخطأ المعاينة. ويعبر عنه بالمعادلة التالية:

$$SS_t = \sum x_t^2 = \sum X_t^2 - \frac{(\sum X_t)^2}{N} \qquad (6.9)$$

وفي مثالنا تكون القيمة

$$SS_t = 6830 - \frac{(444)^2}{30} = 258.8$$

ثم نجد الجزء من المجموع الإجمالي للمربعات الذي يعود إلى انحرافات أوساط المجموعة عن الوسط العام. ويدعى هذا المؤشر (بمجموع المربعات بين المجموعات). ويوجد هذا المؤشر بتطبيق المعادلة التالية:

$$SS_b = \sum x_b^2 = \frac{(\sum X_1)^2}{N_1} + \frac{(\sum X_2)^2}{N_2} + \ldots - \frac{(\sum X)^2}{N} \qquad (6.10)$$

وفي مشكلتنا هذه تكون القيمة

$$SS_b = \sum x_b^2 = \frac{(150)^2}{10} + \frac{(169)^2}{10} + \frac{(125)^2}{10} - \frac{(444)^2}{30} = 97.4$$

ثم نجد الجزء من المجموع الإجمالي للمربعات التي تعود لانحرافات كل درجة فردية عن وسط مجموعتها. ويدعى هذا المؤشر (مجموع المربعات داخل المجموعات) ونحصل عليه بتطبيق معادلة الدرجة الخام لمجموع مربعات الانحرافات لكل مجموعة ثم إجراء الجمع عبر المجموعات :($SS_w = SS_1+SS_2+SS_3$)

$$SS_w = \sum x_w^2 = \sum X_1^2 - \frac{(\sum X_1)^2}{n_1} + \sum X_2^2 - \frac{(\sum X_2)^2}{n_2} + ... \quad (6.11)$$

وفي مشكلتنا تكون هذه القيمة

$$SS_w = \sum x_w^2 = 2310 - \frac{(150)^2}{10} + 2935 - \frac{(169)^2}{10} + 1585 - \frac{(125)^2}{10} = 161.4$$

ويمكن إيجاد مجموع المربعات داخل المجموعات بطرح مجموع المربعات بين المجموعات من المجموع الإجمالي للمربعات والمعبر عنه بما يلي:

$$SS_w = SS_t - SS_b \quad \text{أو} \quad \sum x_w^2 = \sum x_t^2 - \sum x_b^2 \quad (6.12)$$

وفي حالتنا هذه:

$$SS_w = 258.8 - 97.4 = 161.4$$

الجدول (6.5): خلاصة تحليل التباين لثلاث مجموعات

(1) مصدر التباين	(2) *SS*	(3) *df*	(4) *MS*	(5) *F*	(6) مستوى الدلالة
بين المجموعات	97.4	2	48.70	8.14	0.01
داخل المجموعات	161.4	27	5.98		
المجموع الكلي	258.8	29			

الاختبار الفائي "F" للدلالة: يلخص الجدول (6.5) نتائج حساباتنا مع نتائج حسابات أخرى. ويدرج العمود (1) في الجدول المصادر الثلاثة للتباين: تباين بين المجموعات، تباين داخل المجموعات، والتباين الكلي. أما العمود (2) فيحتوي على كل مجموع للمربعات التي قمنا بحسابها. بينما يدرج العمود (3) عدد درجات الحرية المرتبطة مع كل مصدر للتباين. إن عدد درجات الحرية للتباين بين المجموعات يساوي (G-1) حيث (G) هو عدد المجموعات. وفي مثالنا تكون القيمة 3 (2 = 1 -). أما درجات الحرية للتباين داخل المجموعات فهو (... + n_1 - 1 + n_2 - 1) أو

([N-G] أي العدد الإجمالي للدرجات [N] ناقصاً عدد المجموعات [G]). وفي مثالنا تكون هذه القيمة (27=1-10+1-10+1-10) أو (27=3-30).

أما عدد درجات الحرية للمجموع الكلي للتباين فيساوي (N-1)، وفي مثالنا (29=1-30). ويمكن الحصول على هذه القيمة الأخيرة بإضافة درجات الحرية لما بين المجموعات إلى درجات الحرية داخل المجموعات.

عندئذ تكون الخطوة التالية هي حساب تقديري التباين المعروفين بـ "مربع وسط ما بين المجموعات" و "مربع وسط داخل المجموعات". ويمكن الحصول على هذه القيم بتقسيم مجموع المربعات بين المجموعات وداخل المجموعات على درجات الحرية لكل منها. أما القيم الناتجة عن ذلك فهي مربعات الأوساط. وفي مثالنا يكون مربع الوسط بين المجموعات هو (48.7 = 2 / 97.4) أما مربع الوسط داخل المجموعات فهو (5.98 = 27 / 161.4). إن مربع الوسط داخل المجموعات فهو حد الخطأ للنسبة F. وبتطبيق المعادلة التالية نصل أخيراً إلى الناتج النهائي لإجراء تحليل التباين. أي النسبة F.

$$F=\frac{MS_b}{MS_w}=\frac{SS_b/df_b}{SS_w/df_w} \qquad (6.13)$$

وفي مثالنا تكون قيمة

$$F=\frac{48.70}{5.98}=8.14$$

ونعود الآن إلى الجدول A.3 في الملحق لتحديد ما إذا كان النسبة الفائية (F-ratio) دالة إحصائياً. ونجد العمود المعنون درجات الحرية لما بين المجموعات (البسط) ونهبط في العمود إلى مدخل الصف المتطابق مع عدد درجات الحرية لداخل المجموعات (المقام). وعند هذه النقطة في العمود نجد قيمتين، واحدة بنوع فاتح وواحدة بنوع غامق. فإذا كانت النسبة الفائية تساوي أو أكبر من القيمة المعطاة في الفاتح فإنها ذات دلالة على مستوى (.05). وإذا كانت النسبة الفائية التي حصلنا عليها مساوية أو أكبر من القيمة المعطاة في النوع الغامق فإنها دالة أيضاً على مستوى (.01). وفي مثالنا مع درجات الحرية (2) و (27) نحتاج إلى نسبة فائية مقدارها (3.35) لرفض الفرضية الصفرية على مستوى (.05) ونسبة فائية مقدارها (5.49) لرفض الفرضية الصفرية على مستوى (.01). ولأن النسبة الفائية التي حصلنا عليها هي أكبر من كلا هاتين القيمتين فإنها دالة على مستوى (.01) وترفض الفرضية الصفرية على هذا المستوى.

إن الافتراض الذي يشكل الأساس لأسلوب تحليل التباين يشير إلى أنه إذا كانت المجموعات التي ستجري مقارنتها، عينات عشوائية حقاً ومن نفس المجتمع الإحصائي فإن مربع وسط ما بين المجموعات لا ينبغي أن يختلف عن مربع وسط داخل المجموعات بأكثر من المقدار

الذي نتوقعه من الصدفة وحدها. وهكذا، فتحت فرضية صفرية حقيقية نتوقع أن تكون النسبة الفائية تساوي (1.0) تقريباً. ومن ناحية أخرى، إذا كانت الفرضية الصفرية كاذبة فالفرق بين أوساط المجموعات سيكون أكبر مما هو متوقع عن طريق الصدفة، لذا فإن مربع الوسط البيني سيكون أكبر من مربع الوسط الداخلي. وفي هذه الحالات، فإن النسبة F وهي حاصل قسمة مربع الوسط البيني على مربع الوسط الداخلي، ستكون قيمتها أكبر من (1.0). ونلجأ بعدئذ إلى جدول القيم الفائية (A.3) لتحديد ما إذا كانت نسبة البيانات أكبر من (1.0) كافية بما يمكننا من رفض الفرضية الصفرية على مستوى محدد مسبقاً. وكلما أزداد الفرق بين مربعات الأوساط هذه، ازدادت النسبة الفائية وقل احتمال كون الفرضية الصفرية صحيحاً.

عندما ترفض الفرضية الصفرية كنتيجة لإجراء تحليل التباين هذا، فلا يسعنا القول اكثر من أن القياسات التي تم الحصول عليها من المجموعات المعنية تختلف، وأن الفروقات أكبر مما يتوقع المرء وجوده بالصدفة وحدها.

والنسبة الفائية الدال لا تعني بالضرورة أن جميع المجموعات تختلف بصورة دالة عن المجموعات الأخرى. لذا فإن F الدالة قد تكون نتيجة فرق موجود بين مجموعة واحدة وبقية المجموعات. ففي مشكلتنا، مثلاً، قد تكون المجموعة (3) مختلفة بصورة دالة عن المجموعتين (1) و (2)، إلا إن المجموعتين (1) و (2) لا تختلفان بصورة دالة عن بعضيهما. وهناك عدة اختبارات إحصائية يمكن تطبيقها لإيجاد موقع الفروقات الدالة. ويصف كتاب (Glass & Hopkins, 1984) العديد من هذه النصوص وتطبيقاتها.

لقد انتخبنا في مثالنا، المجموعات الثلاث بصورة عشوائية من نفس المجتمع الإحصائي، لذا يوسعنا أن نفترض أنها لم تختلف بما هو أبعد من توقع الصدفة قبيل معالجتنا التجريبية. وتوضح النسبة الفائية أن الفروقات الموجودة بين هذه المجموعات، بعد المعالجة، تفوق توقع الصدفة. ونحن نعزو ذلك مؤقتاً / تجريبياً إلى المعالجة التجريبية، ونستنتج أن مستوى التوتر يؤثر على أداء الأفراد في واجبات حل المشكلة البسيطة. وهذا هو أبعد ما يسعنا الذهاب إليه في تفسيرنا للنسبة الفائية هذه. وإذا احتجنا إلى تحليل إحصائي آخر، فبوسعنا أن نستخدم اختبارات أخرى لتحديد دلالة الفرق بين مجموعات محددة أو مجموعة مؤتلفة من مجموعات. لنفترض أننا خصصنا قبل التجربة مجموعة اللاتوتر كمجموعة ضابطة، وخططنا اتباع الاختبار الفائي الذي يختبر دلالة الفروق بين المجموعة الضابطة والمجموعتين الأخريين. ويوضح كتاب Glass & Hopkins أن طريقة دونيت (Dunnett) مناسبة لهذه المقارنات. فحجم تأثير الفرق بين مجموعتي التوتر المعتدل وعدم التوتر هو 1.56 = 11.60 /(12.5-15.0) (المعادلة 5.16). ويوضح اختبار دونيت أن ذلك غير دال على المستوى (.05) ذي الطرفين. فحجم التأثير لمجموعة التوتر المعتدل إزاء مجموعة اللاتوتر هو 2.75 = 1.60/(12.5-16.9). ويشير اختبار دونيت إلى أن ذلك دال على مستوى (.05) ذي الطرفين.

التحليل متعدد العوامل للتباين

Multifactor Analysis of Variance

قد نود الاستقصاء عن الأثر المدمج لمستوى التوتر وحاجة التحصيل / الإنجاز للأداء في واجب حل مشكلة. ولبحث هذه المشكلة سوف نعمل على تغيير كل من مستوى التوتر وحاجة الإنجاز. ويدعى تصميم التجربة لاستقصاء الآثار المدمجة لمتغيرين مستقلين أو أكثر (بالتصميم العاملي) ويتم تحليل النتائج بواسطة "التحليل متعدد العوامل للتباين".

دعنا نفترض أننا قمنا بهذه التجربة باستخدام خمسة أفراد في كل مجموعة، وأن البيانات الموضحة في الجدول (6.6) تبين خلاصة لملاحظاتنا عن أداء الأفراد. إن تطبيق التحليل متعدد العوامل للتباين سوف يمكننا من معرفة (1) ما إذا كان هناك فرق دال بين أداء جميع الأفراد تحت ظرف توتر عال وأداء جميع من كانوا تحت ظرف توتر متدن، (2) وما إذا كان هناك فرق دال بين أداء جميع الأفراد الذين هم بحاجة عالية للإنجاز وجميع الأفراد ذوي حاجة الإنجاز المتدنية (3) وما إذا كانت اندماجات / توفيقات متغايرة للتوتر وحاجة الإنجاز تؤثر على أداء الأفراد. أما الآثار المستقصاة عن طريق التحليلين الأول والثاني فتدعى "الآثار الرئيسية"، بينما يشار إلى الثالث "بالأثر التفاعلي". والنواتج النهائية لهذه التحليلات ستكون ثلاث نسب فائية يوضح اثنان منها دلالة التأثيرين الرئيسين والثالث يوضح الأثر التفاعلي.

ويشمل حساب هذه النسب الفائية الخطوات التالية:

1- إيجاد المجموع الإجمالي للمربعات، ومجموع المربعات بين المجموعات، ومجموع المربعات داخل المجموعات، باستخدام الإجراءات ذاتها والمعادلات المطبقة في التحليل البسيط للتباين. وهذه القيم المستمدة من البيانات في الجدول (6.6) هي كما يلي:

الجدول (6.6): قياسات واجبات حل المشكلات لأفراد ذوي حاجة إنجاز متدنية وعالية تحت ظروف توتر عالية ومتدنية

حاجة الإنجاز / التحصيل	التوتر: عالـــي	التوتر: متدنـــي	
عالـــي	مجموعة 1 20 20 19 19 17 ΣX 95 $\bar{X} = 19$	مجموعة 3 23 22 21 20 19 ΣX 105 $\bar{X} = 21$	$\Sigma X_{r_1} = 200$ $\bar{X}_{r_1} = 20.0$
متدني	مجموعة 2 22 21 20 19 18 ΣX 100 $\bar{X} = 20$	مجموعة 4 18 16 15 14 12 ΣX 75 $\bar{X} = 15$	$\Sigma X_{r_2} = 175$ $X_{2} = 17.5$
	$\Sigma X_{c_1} = 195$ $\bar{X}_{c_1} = 19.5$	$\Sigma X_{c_2} = 180$ $\bar{X}_{c_2} = 18.0$	ΣX (الكل) $= 375$ $\bar{\bar{X}}$ (الوسط العام) $= 18.75$

$$SS_t = 7181 - \frac{(375)^2}{20} = 149.75$$

$$SS_b = \frac{(95)^2}{5} + \frac{(105)^2}{5} + \frac{(100)^2}{5} + \frac{(75)^2}{5} - \frac{(375)^2}{20} = 103.75$$

$$SS_W = 149.75 - 103.75 = 46.00$$

2- تقسيم مجموعة المربعات بين المجموعات إلى ثلاثة مجاميع مربعات منفصلة - (أ) مجموع المربعات بين الأعمدة، (ب) مجموع المربعات بين الصفوف، (جـ) مجموع مربعات التفاعل بين الأعمدة والصفوف كما مبين أدناه:

أ. يمثل مجموع المربعات بين الأعمدة مجموع مربعات الانحرافات الناشئة عن الفروق بين أوساط الأعمدة والوسط العام. ويمكن إيجاد ذلك باستخدام المعادلة (6.14).

$$SS_{bc} = \frac{(\Sigma X_{c1})^2}{n_{c1}} + \frac{(\Sigma X_{c2})^2}{n_{c2}} + \ldots - \frac{(\Sigma X)^2}{N} \qquad (6.14)$$

وباستخدام هذه المعادلة يكون مجموع المربعات بين الأعمدة للبيانات الموضحة في الجدول (6.6) هو:

$$SS_{bc} = \frac{(195)^2}{10} + \frac{(180)^2}{10} - \frac{(375)^2}{20} = 11.25$$

ب. إن مجموع المربعات بين الصفوف هو مجموع مربعات الانحرافات الناشئة عن الفروق بين أوساط الصفوف والوسط العام. وهو يتم إيجاده بتطبيق المعادلة:

$$SS_{br} = \frac{(\sum X_{r1})^2}{n_{r1}} + \frac{(\sum X_{r2})^2}{n_{r2}} + \ldots - \frac{(\sum X)^2}{N} \qquad (6.15)$$

وبالنسبة للبيانات المقدمة في الجدول (6.6) تكون هذه القيمة:

$$SS_{br} = \frac{(200)^2}{10} + \frac{(175)^2}{10} - \frac{(375)^2}{20} = 31.25$$

جـ. إن مجموع المربعات للتفاعل هو ذلك الجزء من الانحراف بين أوساط المجموعات والوسط الكلي الذي لا يعزى إلى فروقات الصفوف ولا إلى فروقات الأعمدة. وبالتعبير عن ذلك في معادلة يصبح:

$$SS_{int} = SS_b - (SS_{bc} + SS_{br}) \qquad (6.16)$$

وبالكلمات، فإن مجموع مربعات التفاعل يساوي مجموع المربعات بين المجموعات ناقصاً مجموع المربعات بين الأعمدة ومجموع المربعات بين الصفوف:

وبالنسبة للبيانات المقدمة في الجدول (6.6) فإن مجموع مربعات التفاعل هذا يكون:

$$SS_{int} = 103.75 - (11.25 + 31.25) = 61.25$$

3- تحديد درجات الحرية المرتبطة مع كل مصدر تباين. ويمكن إيجادها كما يلي:

df لمجموع المربعات بين الأعمدة(*) = C-1.

df لمجموع المربعات بين الصفوف = R-1.

df للتفاعل = (C-1)(R-1).

df لمجموع المربعات بين المجموعات = G-1.

df لمجموع المربعات داخل المجموعات = N-G.

df لمجموع المربعات الكلي = N-1.

(*) يشار أحياناً إلى مجموع المربعات بين الأعمدة بصيغة أخرى، هي "مجموع مربعات الأعمدة البينية" ويشار إلى "بين الأعمدة" بصيغة "الأعمدة البينية" و "بين الصفوف" بصيغة "الصفوف البينية"، و "داخل المجموعات" بصيغة "المجموعات الداخلية" ونفضّل استخدام الصيغة الأساسية إذ أنها أوضح. (المراجع)

حيث:

C = عدد الأعمدة.

R = عدد الصفوف.

G = عدد المجموعات.

N = عدد الأفراد في كل المجموعات.

4- إيجاد قيم مربع الوسط بتقسيم كل حاصل جمع للمربعات على عدد درجات الحرية لكل منها.

5- حساب النسبة F للآثار الرئيسية والتفاعلية وذلك بتقسيم مجموع مربعات أوساط المجموعات البينية على مربع وسط المجموعات الداخلية لكل من العناصر الثلاثة.

إن نتائج الحسابات القائمة على أساس البيانات المقدمة في الجدول (6.6) ملخصة في الجدول (6.7) – وتدرج ثلاث نسب فائية في هذا الجدول. ولإيجاد دلالة كل من هذه القيم نعود إلى الجدول A.3 (جدول قيم F) كالسابق. وللدخول نستخدم عدد درجات الحرية المرتبطة مع كل نسبة فائية (df للبسط) وعدد درجات الحرية المرتبطة مع مربع الوسط لداخل المجموعات (df للمقام). فمثلاً، إن النسبة الفائية للأعمدة البينية "بين الأعمدة" هو (3.913). وبالرجوع إلى الجدول نرى أنه مع درجات حرية 1 و 16، هناك حاجة لنسبة فائية مقدارها (4.49) أو أكثر للدلالة على مستوى (.05). ولأن النسبة الفائية أصغر من القيمة المبينة في الجدول، نستنتج أن فرق التوتر العالي إزاء فرق التوتر المنخفض غير دال إحصائياً.

ونظراً لأن درجات الحرية للصفوف البينية "بين الصفوف" هو نفس درجات الحرية بين الأعمدة فإن نفس المستويات لـ F، 4.49 (مستوى .05) و (8.53) (مستوى .01) يمكن تطبيقها. ولأن القيمة الناتجة لـ F، (10.869) تفوق كلا من هاتين القيمتين فإنها دالة على مستوى (.01).

الجدول (6.7): خلاصة التحليل متعدد العوامل للتباين، 2×2

مصدر التباين	SS	df	MS	F	مستوى الدلالة
بين الأعمدة (التوتر)	11.25	1	11.25	3.913	--
بين الصفوف (حاجة الإنجاز)	31.25	1	31.25	10.869	.01
التفاعل بين الصفوف والأعمدة	61.25	1	61.25	21.304	.01
بين المجموعات	103.75	3	34.583		
داخل المجموعات	46.00	16	2.875		
المجموع	149.75	19			

وفي مثالنا، فإن درجات الحرية للتفاعل هي نفسها (1، 16). وعليه هناك حاجة لنفس قيم F . فالقيمة الفائية (21.304) أكبر من (8.35)، وعليه فإن التفاعل دال على مستوى (01.). أما الفرضية الصفرية – أي الفرق بين وسطي مجموعتي التوتر العالي والتوتر المنخفض بين أفراد الحاجة المنخفضة للإنجاز هو نفسه كالفرق بين وسطي مجموعتي التوتر العالي والتوتر المنخفض بين أفراد الحاجة العالية للإنجاز – يمكن رفضها.

تفسير النسبة الفائية Interpretation of The F-Ratios

لم تكن النسبة الفائية الأولى (بين الأعمدة) في الجدول 6.7 دالة وهو يبين أن ظروف التوتر لا تختلف في تأثيرها بشكل دال فيما بينها على أداء الأفراد في التجربة. وهذا التحليل هو مقارنة الأداء المدمج للمجموعتين 2, 1 ($\overline{X}=19.5$) مع الأداء المدمج للمجموعتين 4, 3 ($\overline{X}=18.0$). وكان بوسعنا التوصل إلى نفس الاستنتاج باستخدام الاختبار التائي (t-test).

أما النسبة الفائية الثانية (بين الصفوف) التي كانت دالة على مستوى (01.) فهي تستند على مقارنة أداء الأفراد في المجموعتين 3, 1 ($\overline{X}=20.0$)المجموعتين 4, 2 ($\overline{X}=17.5$). ومن دلالة هذه النسبة الفائية يمكن أن نستنتج أن الفرق بين أداء الأفراد ذوي الحاجة للإنجاز العالي، وأولئك ذوي الحاجة للإنجاز المنخفض، هي أبعد من توقع الصدفة. ولأن لدينا نسبة فائية دالة للفرق، فنستنتج أنه تحت ظروف مشابهة لتلك التي في تجربتنا، فإن مستوى عال لأداء الواجب يمكن توقعه من أشخاص ذوي حاجة إنجاز عالية.

وتبين النسبة الفائية الثالثة أثر التفاعل بين متغيرين: مستوى التوتر وحاجة الإنجاز. أما دلالة النسبة الفائية في هذه الحالة، فتعني أن أثر مستوى التوتر على الأداء في واجب حل المشكلة يعتمد على درجة الحاجة للإنجاز. ويمكن أن نرى هذه الظاهرة بشكل أوضح لو قارنا النتائج الملاحظة مع النتائج المتوقعة عندما لم يكن هناك تفاعل بين المتغيرين المستقلين.

دعنا نحسب ما قد نتوقعه حول أوساط المجموعات الأربع في حالة عدم وجود تفاعل. إن وسط جميع الأفراد (18.75). أما وسط عشرة أفراد تحت توتر عال (19.5) هو (0.75) أكبر من هذا العدد، بينما وسط عشرة أفراد تحت توتر منخفض أقل بمقدار 0.75. فالوسط للعشرة أفراد ذوي الحاجة العالية للإنجاز (20) هو (1.25) أكبر من وسط جميع الأفراد، لكن وسط عشرة أفراد ذوي حاجة لإنجاز منخفض هو (0 .75) أقل.

ويمكن أن نحسب لكل مجموعة الوسط المتوقع لهذه المجموعة في حال عدم وجود تفاعل. ونقوم بذلك بالإضافة إلى الوسط العام فرق العمود الذي تكون المجموعة فيه وفرق الصف الذي تكون المجموعة فيه. وإذا لم يكن هناك تفاعل فماذا يمكن أن نتوقع أن يكون وسط المجموعة 1؟ فإذا بدأنا بالوسط الإجمالي (18.75) فسوف نضيف، (0.75) لأن الأفراد تحت توتر عال وآخر (1.25) لأنه كانت لديهم حاجة إنجاز عالية. وهذا يعطينا إجمالياً مقداره (20.75). وبمتابعة هذا الأجراء لكل من المجموعات الأربع نحصل على القيم المتوقعة التالية:

	الوسط الإجمالي +	فرق التوتر +	فرق الحاجة للإنجاز =	القيمة المتوقعة
المجموعة 1	18.75	0.75+	1.25+	20.75
المجموعة 2	18.75	0.75+	1.25-	18.25
المجموعة 3	18.75	0.75-	1.25+	18.25
المجموعة 4	18.75	0.75-	1.25-	18.25

والآن، قارن أوساط المجموعات الفعلية مع أوساط المجموعات المتوقعة:

المتوقع

	عالي	منخفض	
عالي	مجموعة 1 $\overline{X}$=20.75	مجموعة 3 $\overline{X}$=19.25	$\overline{X}$=20
منخفض	مجموعة 2 $\overline{X}$=18.25	مجموعة 4 $\overline{X}$=16.75	$\overline{X}$=17.5
	$\overline{X}$=19.5	$\overline{X}$=18.0	$\overline{\overline{X}}$=18.75

الفعلي

	عالي	منخفض	
عالي	مجموعة 1 $\overline{X}$=19	مجموعة 1 $\overline{X}$=21	$\overline{X}$=20
منخفض	مجموعة 2 $\overline{X}$=20	مجموعة 4 $\overline{X}$=15	$\overline{X}$=17.5
	$\overline{X}$=19.5	$\overline{X}$=18.0	$\overline{\overline{X}}$=18.75

(لاحظ أن بوسعنا استخدام الفروقات بين القيم المتوقعة والقيم الملاحظة لحساب مجموع المربعات للتفاعل مباشرة. فكل مجموعة تختلف عن الوسط المتوقع بـ (1.75)، قم بتربيع هذه القيمة وأضرب بعدد الحالات لتحصل على $1.75^2 \times 20 = 61.25$).

نرى أن المجموعة (1) قامت بأقل ما نتوقع حيث نعرف أن أفرادها كانوا تحت توتر عال ولهم حاجة إنجاز عالية، أما المجموعة (2) التي لها حاجة إنجاز متدنية وواقعة تحت توتر عال فقد قامت بأفضل مما نتوقعه. وإذا أخذنا بنظر الاعتبار المجموعات الواقعة تحت توتر متدن لوجدنا أن المجموعة (3) ذات الحاجة للأداء العالي قد قامت بأفضل ما هو متوقع، بينما المجموعة (4) ذات الحاجة المنخفضة للأداء فقد قامت بأقل مما نتوقع. ونظراً لأن اختبارنا الفائي أوضح أن التفاعل كان دالاً فإننا نستنتج أن التوتر العالي ينتج درجات أعلى حين يندمج مع حاجة الأداء المنخفضة مما لو اندمج مع حاجة الإنجاز العالية، بينما ينتج التوتر المتدني درجات أعلى عندما يندمج مع حاجة الأداء العالية مما لو اندمج مع حاجة الإنجاز المنخفضة.

ويعتبر استخدام تحليل العوامل المتعددة ذا قيمة كبيرة في البحوث التربوية لأن العديد من المسائل التي يحتاج التربويون إلى استقصائها هي مسائل معقدة في طبيعتها. وتمكننا هذه الأساليب من تحليل الآثار المدمجة لمتغيرين مستقلين أو أكثر فيما يتعلق بمتغير تابع. فمثلاً، إن

مقارنة بسيطة لوسطي المتغير التابع لمجموعتين من الطلبة ثم تدريسهم بطرق مختلفة قد يؤدي إلى نتائج غير دالة. لكنه إذا دمج الذكاء في التجربة كمتغير مستقل مقاس، فقد نجد أن طريقة واحدة قد تعمل أفضل مع الطلبة الأقل ذكاءً بينما تعمل الأخرى بشكل أفضل مع الطلبة الأكثر ذكاءً. (هناك مادة أكثر حول تفسير التحليل العاملي للتباين في الفصل 9).

إن تحليل العوامل متعددة التباين غير محدود بمتغيرين مستقلين كما في مثالنا. وقد يدمج أي عدد من المتغيرات المستقلة في هذا الأسلوب. وتوضح العديد من كتب الإحصاء المتوسطة بما في ذلك كتاب (Glass & Hopkins, 1984) حساب وتفسير هذه الإجراءات.

اختبار مربع كاي للدلالة
THE CHI-SQUARE TEST OF SIGNICICANE

نحتاج أحياناً إلى إيجاد دلالة الفروقات بين أجزاء من الأفراد والأشياء والأحداث، وما إلى ذلك، تقع في أصناف مختلفة. ويدعى الاختبار الإحصائي المستخدم في هذه الحالات باختبار مربع كاي (X^2).

في اختبار مربع كاي تتم مقارنة مجموعتين من التكرارات: التكرارات الملاحظة والتكرارات المتوقعة، أما التكرارات الملاحظة، كما يدل أسمها، فهي التكرارات الفعلية التي نحصل عليها عن طريق الملاحظة، بينما التكرارات المتوقعة هي تكرارات نظرية تستخدم لغرض المقارنة.

مربع كاي لمتغير واحد (حسن المطابقة)
The One-Variable Chi-Square (Goodness of Fit)

لننظر في الفرضية التي تقول إن أجزاء الطلبة الإناث والذكور في مقررات مادة الإحصاء تختلف عن أجزاء الطلبة من الإناث والذكور في كلية التربية ككل. وإذا عرفنا أن 40 في المائة من مجموع المسجلين في الكلية هم من الذكور وأن 300 من الطلبة سجلوا في مقررات الإحصاء فإن التكرارات المتوقعة للطلاب والطالبات المسجلين في الإحصاء ستكون:

طلبة ذكور	120	300
طالبات	180	

الآن افترض أن التكرارات الملاحظة كانت

طلاب(*)	140	300
طالبات	160	

(*) هناك عرف سائد باستخدام "طلبة" للذكور والإناث بينما "طلاب" تشير إلى الذكور . (المراجع)

ونحن نريد أن نحدد ما إذا كان الفرق بين التكرارات الملاحظة والمتوقعة دالا إحصائياً. ولتحديد ذلك نطبق معادلة مربع كاي التي هي (6.17):

$$X^2 = \sum \left[\frac{(f_o - f_e)^2}{f_e} \right] \qquad (6.17)$$

حيث:

X^2 = قيمة مربع كاي.
f_o = التكرار الملاحظ.
f_e = التكرار المتوقع.

وبتطبيق هذه المعادلة على البيانات نحصل على:

$$X^2 = \frac{(140-120)^2}{120} + \frac{(160-180)^2}{180} = 5.55$$

ولتحديد ما إذا كانت قيمة مربع كاي دالة، نرجع إلى جدول قيم X^2 في الملحق (الجدول A.4). العمود الأول في الجدول يبين عدد درجات الحرية في أي مسألة لمربع كاي. وتقدم بقية الأعمدة القيم المطلوبة لمستويات مختلفة من الدلالة. أما عدد درجات الحرية، كما ناقشنا سابقاً، فتستند إلى عدد الملاحظات الحرة في التغير حينما توضع بعض القيود على البيانات.

الجدول (6.8)
التكرارات الملاحظة لإجابات 120 طالب كلية
حول ما يفضلونه بشأن أماكن سكنهم

الأفراد	القسم الداخلي	المدينة	عدم تفضيل	المجموع
طلبة كلية	40	50	30	120

الجدول (6.9)
التكرارات المتوقعة لإجابات 120 طالب كلية
حول ما يفضلونه بشأن أماكن سكنهم

الأفراد	القسم الداخلي	المدينة	عدم تفضيل	المجموع
طلبة كلية	40	40	40	120

وحينما يكون لدينا عدد محدد من الملاحظات مقسمة في صنفين فقط، وحينما يتحدد العدد الواقع في صنف معين، يثبت العدد الآخر. وهكذا، فعندما نجد أن عدد الطلبة الذكور هو (140) فإن عدد الطلبة الإناث من مجموع الـ (300) لابد أن يكون (160). وفي هذا المثال هناك درجة حرية واحدة. وفي مسائل كهذه يكون عدد درجات الحرية مساوياً (K-1) حيث يعتبر (K) عدد الأصناف المستخدمة في التصنيف. وبالرجوع إلى جدول قيم X^2 نجد أن قيمة الملاحظة 5.55 دالة إحصائيا على مستوى (.05) و (.02).

وبتفسير هذه النتيجة يمكن أن نشير الآن إلى أن نسبة الذكور الذين أخذوا مقررات في الإحصاء هي أكبر من النسبة في كلية التربية برمتها على مستوى (.05) من الثقة. إن مستوى الدلالة (.05) يعني أن هناك أقل من خمس فرص في كل مائة ملاحظة لمثل هذا الفرق بين نسب الطلبة الذكور والإناث من خلال الصدفة وحدها، وهكذا تدعم البيانات فرضية البحث بأن نسب الذكور والإناث من الطلبة ممن أخذوا مقررات في الإحصاء مختلفة بشكل دال عن نسبة الذكور والإناث في مجمل كلية التربية .

إن استخدام اختبار مربع كاي غير محدد بالأوضاع التي يكون فيها صنفان فقط للتصنيف، إذ يمكن استخدام هذا الاختبار لاختبار الفرضية الصفرية التي تنص على عدم وجود فرق دال بين أجزاء الأفراد الذين يقعون في عدد من الأصناف المختلفة. أفترض، مثلاً أننا طلبنا من عينة من (120) طالب كلية ما إذا كانوا يفضلون العيش في قسم داخلي أو في المدينة أو ليس لديهم أي تفضيل مع النتائج المبينة في الجدول (6.8).

فإذا لم يكن هناك فرق بين أصناف الإجابة فسوف يكون لدينا (40) إجابة في كل معيار. هذه ستكون تكراراتنا المتوقعة كما هي مبينة في الجدول (6.9).

إن مقارنة بمجموعتي التكرارات المقدمة في الجدولين (6.8) و (6.9) تبين أن هناك فروقاً بين بياناتنا الملحوظة والمتوقعة. ولتحديد ما إذا كانت دالة، نطبق اختبار مربع كاي. وتكون قيمة X^2 لهذه البيانات باستخدام المعادلة (6.17).

$$X^2 = \frac{(40-40)^2}{40} + \frac{(50-40)^2}{40} + \frac{(30-40)^2}{40} = 5.00$$

إن درجات الحرية، مرة أخرى، يساوي عدد الأصناف ناقصاً (1) أي (K-1)، أو في هذه الحالة (2 = 3-1). وبالعودة إلى جدول قيم (X^2) نرى أنه مع درجتين من الحرية تكون قيمة (X^2) التي نحتاجها لـ (5.991) أو أكبر للدلالة على مستوى (.05). وعلى أية حال، فإن قيمة X^2 الناتجة أصغر من هذه القيمة وعليه لا تكون دالة إحصائياً. ويعني هذا أن الفروقات الملاحظة بين الأصناف قد حدثت بلا ريب عن طريق الصدفة. وعليه فإن الفرضية الصفرية – أي عدم وجود فرق دال بين تكرارات الأصناف الثلاثة – لا يمكن رفضها. وبعبارة أخرى إذا كانت نسب

التفضيلات للأصناف الثلاثة في مجتمع طلبة الكلية برمته متساوية فنتوقع أن نلاحظ فروقات العينة بنفس حجم الفروقات في عينتنا أكثر من خمس مرات من مائة فرصة.

مربع كاي ذي المتغيرين (اختبار الاستقلال)
The Two-Variable Chi-Square (Independence Test)

كنا قد درسنا حتى هذا الحد أمثلة صنفت فيها الملاحظات في بعد واحد. ونرغب أحياناً، على أية حال، في استخدام أكثر من بعد للتصنيف. أفترض مثلاً أننا أضفنا بعداً آخر إلى المسألة السابقة وطلبنا من طلبة الكلية والخريجين إيضاح تفضيلاتهم حول أماكن سكنهم. أفترض التكرارات المبينة في الجدول (6.10) كانت هي النتيجة.

جدول (6.10): التكرارات الملاحظة لإجابات 200 طالب كلية وطالب خريج بشأن تفضيلهم أماكن سكنهم

الأفـــراد	القسم الداخلي	المدينة	عدم تفضيل	المجموع
طلبة كلية	40	50	30	120
خريجون	20	40	20	80
المجموع	60	90	50	200

في هذه الحالة تكون الفرضية الصفرية هي أن تفضيل أماكن العيش هي نفسها للخريجين ولطلبة الكلية – أي لا توجد علاقة بين المتغيرين (حالة الطالب) و (تفضيل مكان السكن) أو هما مستقلان. وتكون الفرضية الصفرية في اختبار مربع كاي للاستقلال دوما هي أن المتغيرات مستقلة في المجتمع الإحصائي. وتبين ملاحظاتنا أن 30 في المائة من مجموع الطلبة يفضلون الأقسام الداخلية، 45 في المائة يفضلون المدينة و 25 في المائة لم يبدو أي تفضيل.

فإذا كانت الفرضية الصفرية صحيحة فنتوقع أن نجد نفس هذه النسب بين طلبة الكلية والخريجين. ويمكننا حساب التكرارات المتوقعة وذلك بضرب تكرار الصف المرتبط مع الخلية بتكرار العمود المرتبط مع تلك الخلية ثم تقسيم هذا الناتج على المجموع الأكبر ($E = f_r f_c / N$). فمثلاً، التكرار المتوقع لإجابة طلبة الكلية ممن يريدون العيش في القسم الداخلي هو ($120 \times 60 \div 200 = 36$)، أما بالنسبة لطلبة الكلية الذين يفضلون العيش في المدينة فيكون ($120 \times 90 \div 200 = 54$)، والطلبة الخريجون الذين يريدون العيش في قسم داخلي هو ($80 \times 60 \div 200 = 24$)، وباستخدام هذه الطريقة نجد التكرارات المتوقعة لكل خلية كما مبينة في الجدول (6.11).

جدول (6.11): التكرارات المتوقعة لإجابات 200 طالب كلية وخريج حول تفضيلهم فيما يخص أماكن سكنهم

الأفراد	القسم الداخلي	المدينة	عدم تفضيل	المجموع
طلبة كلية	36	54	30	120
خريجون	24	36	20	80
المجموع	60	90	50	200

لاحظ أن جميع نواتج الجمع الصفي والعمودي في الجدول (6.11) تطابق نواتج الجمع المبينة في الجدول (6.10). ونتساءل الآن ما إذا كانت التكرارات الملاحظة تختلف بما يكفي عن التكرارات المتوقعة بما يمكننا رفض الاحتمال القائل إن هذه الفروقات يمكن لها أن تحدث بمجرد الصدفة. وبتطبيق المعادلة نحصل على

$$X^2 = \frac{(40-36)^2}{36} + \frac{(50-54)^2}{54} + \frac{(30-30)^2}{30} + \frac{(20-24)^2}{24} + \frac{(40-36)^2}{36} + \frac{(20-20)^2}{20}$$

$$X^2 = 1.8518$$

ونحصل على عدد درجات الجدول ذي الطريقتين (المتغيرين) بتطبيق المعادلة:

$$df = (C\text{-}1)(R\text{-}1) \qquad (6.18)$$

حيث:

df = عدد درجات الحرية.

C = عدد الأعمدة.

R = عدد الصفوف

وبتطبيق هذه المعادلة على المسألة قيد الدرس نحصل

$$df = (3\text{-}1)(2\text{-}1) = 2$$

وبالعودة إلى الجدول A.4 نرى أنه مع درجتي حرية تكون هناك حاجة لـ (X^2) قيمة تساوي (5.991) للدلالة على مستوى (.05). إلا أن قيمة (X^2) وهي (1.8518) أصغر من هذه القيمة المجدولة، وعليه فهي غير دالة، ويعني ذلك أن الفروقات بين التكرارات المتوقعة والملاحظة لا تتجاوز ما هو متوقع بالصدفة. وبعبارة أخرى ليس لدينا دليل ثقة بأن هناك علاقة بين المتغيرات (حالة الطلبة) و (تفضيل مكان العيش) في المجتمع الإحصائي الذي نستمد منه عينتنا. ومن ثم لا يمكن رفض الفرضية الصفرية.

افتراضات مربع كاي Assumptions of Chi-Square

إن سهولة استخدام مربع كاي، قد تجعل المرء ينسى أن هناك افتراضات يجب تلبيتها لتقديم تفسيرات صادقة.

1- الملاحظات يجب أن تكون مستقلة – أي يجب أن يكون الأفراد في كل عينة قد تم انتخابهم بشكل عشوائي ومستقل.

2- يجب أن تكون الأصناف مقتصرة بشكل متبادل، فبوسع كل ملاحظة أن تظهر مرة واحدة وواحدة فقط في الصنف العائد للجدول.

3- تقاس الملاحظات كتكرارات.

الخلاصة SUMMARY

يأمل الباحثون في صياغة تعميمات حول المجتمعات الإحصائية وذلك عن طريق دراسة مجموعات من الأشخاص المنتخبين من مجتمعات إحصائية. وستكون هذه التعميمات معقولة فقط، إذا كانت المجموعات المنتخبة – العينات – المستخدمة في هذه الدراسات ممثلة لمجموعات كبيرة – مجتمعات إحصائية – التي تنتخب منها.

يتم التمييز بين نوعين كبيرين من إجراءات المعاينة: المعاينة الاحتمالية والمعاينة اللااحتمالية. وتتسم المعاينة الاحتمالية باستخدام الاختيار العشوائي لعناصر المجتمع الإحصائي. أما في المعاينة اللااحتمالية، فإن حكم الباحث يحل محل الانتخاب العشوائي. إن المعاينة العشوائية البسيطة والمعاينة الطبقية والمعاينة العنقودية والمعاينة المنتظمة هي أشكال من المعاينة الاحتمالية. وفي المعاينة العشوائية البسيطة يكون لجميع أفراد المجتمع الإحصائي فرصة متساوية للانضمام إلى العينة. وفي المعاينة الطبقية يتم انتخاب عينات مستقلة من مجموعات فرعية مختلفة أو أصناف من مجتمع إحصائي. أما في المعاينة العنقودية فإن المجموعات أو العناقيد التي تحدث بصورة طبيعية تنتخب من مجتمع إحصائي، ثم يستخدم جميع الأفراد داخل العناقيد المنتخبة كعينة. وأخيراً، في المعاينة المنتظمة تؤخذ كل حالة Kth من قائمة المجتمع الإحصائي كعينة. أما أشكال المعاينة اللااحتمالية فهي المعاينة العرضية والمعاينة الهادفة وعينة الحصة، وفي المعاينة العرضية تستخدم الحالات المتيسرة كعينة. أما في المعاينة الهادفة، فإن الحالات التي حكم على أنها نموذجية لمجتمع إحصائي تكوّن العينة. وأخيراً، في المعاينة الحصصية يتم تحديد الحصص في أصناف متعددة من المجتمع الإحصائي، ثم تنتخب الحالات التي يحكم عليها بكونها نموذجية لكل صنف.

يقوم الإحصاء الاستدلالي بتزويد وسائل يستطيع الباحثون بواسطتها تقدير مدى ثقتهم في الاستدلال على أن الظواهر التي لوحظت في العينات سوف تلاحظ في مجتمعات إحصائية

استُمدت منها العينات. وبعبارة أخرى، تمكننا الإحصائيات الاستدلالية من تقدير مدى الثقة بملاحظاتنا.

وكاستراتيجية أساسية في الإحصاء الاستدلالي، نقوم بحساب الفرق بين الملاحظات التي يحتمل لها أن تنشأ بالصدفة وحدها. وتدعى نتيجة هذا الحساب "حد الخطأ". ثم تقارن الفروقات الملاحظة بين الملاحظات مع حد الخطأ. فإذا كانت الفروقات الملاحظة مشابهة للفروقات التي قد تنشأ بالصدفة، فلا يمكن للباحث أن يرفض احتمال أن تكون الفروق الملاحظة مجرد عامل صدفة. وإذا كانت الفروقات الملاحظة أكبر من حد الخطأ، فإن الباحث يرجع إلى القيم المجدولة للإحصاءة لتحديد ما إذا كانت نسبة الملاحظة إلى الخطأ كبيرة بما يكفي لرفض تفسير الصدفة على مستوى دلالة محدد مسبقاً.

ويتحدد حجم العينة المطلوب لإعطاء احتمال محدد لرفض الفرضية الصفرية على مستوى محدد من الدلالة وبحجم تأثير معين، باستخدام حسابات القوة.

إن المؤشرات الأكثر شيوعاً في الإحصاء الاستدلالي هي الاختبار التائي، وتحليل التباين واختبار مربع كاي، ويستخدم الاختبار التائي لتحديد ما إذا كان الفرق بين وسطي عينتين دالاً إحصائياً. وهناك ثلاثة أنواع من الاختبارات التائية: (1) الاختبار التائي للمجموعات المستقلة وهو يستخدم لمقارنة وسطي عينتين تستمدان بصورة مستقلة من مجتمع إحصائي، (2) الاختبار التائي للمجموعات غير المستقلة، وهو يستخدم مع عينتين يتم اختيارهما بطريقة المزاوجة أو بقياسين متكررين يتم الحصول عليها من نفس الأفراد و (3) الاختبار التائي لمعامل ارتباط بيرسون.

ويستخدم تحليل التباين لمقارنة وسطي عينتين أو أكثر ولاختبار الفرضية الصفرية التي تقول بعدم وجود فروقات دالة بين أوساط هذه العينات. ويمكننا التحليل متعدد العوامل للتباين من اختبار أثر أكثر من متغير مستقل وأثر تفاعل مثل هذه المتغيرات.

وتعتبر إحصاءة مربع كاي مؤشراً يستخدم لإيجاد دلالة الفروق بين أجزاء الأفراد والأشياء والأحداث وما إلى ذلك والتي تقع في أصناف مختلفة عن طريق التكرارات الملاحظة والتكرارات المتوقعة حسب فرضية صفرية صحيحة.

مفاهيم أساسية Key Concepts

accessible population	مجتمع إحصائي يمكن الوصول إليه، سهل المنال
accidental sampling	معاينة عرضية
analysis of variance (ANOVA)	تحليل التباين (أنوفا)
biased sample	عينة متحيزة
chi-square	مربع كاي

cluster sampling	معاينة عنقودية
degrees of freedom	درجات الحرية
directional test (one-tailed)	اختبار موجه /اتجاهي (بطرف واحد)
expected frequency	تكرار متوقع
factorial design	تصميم عاملي
F-test	اختبار فائي "F"
interaction	تفاعل
level of significance	مستوى الدلالة
main effect	تأثير رئيسي
multifactor analysis of variance	التحليل متعدد العوامل للتباين
nondirectional test (one-tailed)	اختبار غير موجه / اتجاهي (بطرف واحد)
nonprobability sampling	معاينة اللااحتمال (لا احتمالية)
null hypothesis	فرضية صفرية
observed frequency	تكرار ملاحظ
one-tailed test	اختبار بطرف واحد
parameter	معْلم / بارامتر
population	مجتمع إحصائي
power calculation	حساب القوة
probability sampling	معاينة احتمالية
purposive sampling	معاينة هادفة / قصدية
quota sampling	معاينة حصصية
random sample	معاينة عشوائية
sampling error	خطأ المعاينة
standard error of the mean	خطأ معياري للوسط
statistic	إحصاءة
stratified sampling	معاينة طبقية
systematic sampling	معاينة منتظمة
table of random numbers	جدول الأعداد العشوائية
target population	مجتمع إحصائي مستهدف
test of significance	اختبار الدلالة
t-test for independent samples	الاختبار التائي للعينات المستقلة
t-test for nonindependent samples	الاختبار التائي للعينات غير المستقلة

الاختبار التائي لبيرسون	t-test for Pearson r
الاختبار ذو الطرفين	two-tailed test
خطأ النموذج 1	Type I error
خطأ النموذج 2	Type 11 error

تمارين EXERCISES

1. هل إن دقة العينة في تمثيل خواص المجتمع الإحصائي الذي استمدت منه تزداد دائماً مع حجم العينة ؟ أوضح .

2. طُلب منك تحديد ما إذا كان المدرسون في مدرسة المقاطعة المركزية يفضلون مفهوم "المدرسة على مدار السنة". ولأن المقاطعة كبيرة فقد طلبت منك الاتصال فقط بـ 500 مدرس. حدد العدد الذي تختاره من كل من المستويات التالية لسحب عينة عشوائية طبقية نسبية:

المستوى	العدد الإجمالي
ابتدائية	3500
متوسطة	2100
ثانوية	1400
المجموع	**7000**

3. طلب منك القيام بمسح للرأي في كلية ذات مجتمع إحصائي يبلغ 15,000 طالب، كيف ستبدأ في سحب العينة التمثيلية لهؤلاء الطلاب لأجل مسحك؟

4. لدى مجلة محلية مليون مشترك. ويريد موظفو التحرير معرفة أي من جوانب المجلة محببة وأي منها غير محببة. يقرر الموظفون بأن المقابلة الشخصية هي أفضل طريقة للحصول على المعلومات. ولأسباب عملية واقتصادية سيتم استطلاع رأي 500 شخص فقط في خمس مدن. في هذا الوضع، حدد ما يلي:
أ. المجتمع الإحصائي المستهدف .
ب. المجتمع الإحصائي الذي يمكن الوصول إليه / السهل المنال.
جـ. العينة.

5. أي مما يلي هي عينات احتمالية ؟ وأي منها هي عينات لا احتمالية؟
أ. عينة عشوائية

ب. عينة عرضية
جـ. عينة عنقودية
د. عينة طبقية
هـ. عينة هادفة
و. عينة حصصية.
ز. عينة منتظمة.

6. يرغب باحثون دراسة المسألة: هل الشقراوات أكثر فكاهة؟
أ. ما هي الفرضية الصفرية في هذه المسألة.
ب. ما هو خطأ االنوع 1 في هذه الحالة.
جـ. ما هو خطأ النوع 2 في هذه الحالة.
د. إذا استخدم أحد الباحثين مستوى (.05) للدلالة في بحث هذه المسألة وباحث آخر استخدم (.001) مستوى دلالة فأيهم أكثر احتمالاً في صنع خطأ النوع 1؟
هـ. إذا استخدم أحد الباحثين مستوى (.05) من الدلالة في بحث هذه المسألة وباحث آخر استخدم (.001) مستوى دلالة، فأيهم أكثر احتمالاً للوقوع في خطأ النوع 2؟

7. ما الذي بمقدور الإحصائي الاستدلالي تقديمه للباحثين في أي مما يأتي؟
أ. الوصول إلى استنتاجات لا خطأ فيها.
ب. الوصول إلى استنتاجات معقولة بمعلومات غير كاملة.
جـ. إضافة مسحة من الشرعية على ما هو مجرد حدس.

8. ما هما الشرطان الضروريان لحدوث خطأ النوع 1؟

9. أي من العبارات التالية تصف دور الفرضية الصفرية في البحث؟
أ. إنها تمكننا من تحديد احتمال وقوع حادث بالصدفة وحدها عندما لا تكون هناك علاقة حقيقية بين المتغيرات.
ب. تمكننا من البرهنة على وجود علاقة فعلية بين المتغيرات.
جـ. تمكننا من البرهنة على عدم وجود علاقة فعلية بين المتغيرات.

10. أي مما يأتي يوقع المرء في خطأ النوع 2؟
أ. يرفض فرضية صفرية خاطئة.
ب. يرفض فرضية صفرية صحيحة.
جـ. قام بارتكاب خطأ النوع 1.
د. يحتفظ بفرضية صفرية خاطئة.

هـ. يحتفظ بفرضية صفرية صحيحة.

11. إلى أي من الأوضاع التالية تشير عبارة "مستوى الدلالة"؟
أ. احتمال حادث، بسبب الصدفة وحدها، الذي يحسب بعد تحليل بيانات من تجربة.
ب. احتمال خطأ النوع 1، الذي يرغب الباحث في قبوله.
جـ. الاحتمال الفعلي في خطأ النوع 2.
د. احتمال خطأ النوع 2 الذي يرغب الباحث في قبوله.

12. كيف يحدد المرء مستوى الدلالة الذي يستخدمه في تجربة؟

13. استخدمت شركة لتصنيع السكائر باحثين لمقارنة معدل حدوث سرطان الرئة بين المدخنين وغير المدخنين. وبدراسة نتائج البحوث السابقة عن هذه المسألة، فما هو الذي ستحث هذه الشركة الباحثين للحرص عليه، بشكل خاص، لتجنب خطأ النوعين 1، 2؟

14. ماذا يتضح عندما لا تكون نتائج دراسة معينة دالة إحصائياً؟

15. قارن المعاينة الطبقية مع المعاينة الحصصية؟

16. ما هي العناصر الثلاثة التي ينبغي اعتبارها في تحديد حجم العينة؟

17. ما مدى حجم العينة المطلوبة لحجم تأثير (.10) أو أكبر لكي تكون فرضتها 80 في المائة في أن تكون دالة إحصائيا على مستوى دلالة (.05) بطرفين؟

18. جد قيمة مربع كاي للبيانات التالية. ثم وضح ما إذا كانت القيمة الناتجة لمربع كاي دالة على مستوى دلالة (.05).

2	8	10
6	4	10
8	12	

19. لدينا إجابات مجموعتين من الطلبة (تخصص وغير تخصص تربية) على بند في مقياس ليكرت للاتجاهات. جد قيمة مربع كاي لهذه المسألة ووضح ما إذا كانت قيمة مربع كاي الناتجة دالة إحصائياً.

	موافق بقوة	موافق	غير مقرر	غير موافق	غير موافق بقوة
طلبة التربية	7	8	10	9	6
طلبة آخرون	8	10	7	8	7

20. لاختبار كل فرضية أدناه، أوضح أياً من الإجراءات الإحصائية التالية يكون مناسبا؟

- الاختبار التائي للأوساط المستقلة
- الاختبار التائي للأوساط غير المستقلة
- التحليل العاملي للتباين
- مربع كاي

أ. نسبة طلبة الدكتوراه في الصف ممن يحصلون على تقديرات "A" سيكون أكبر من نسبة طلبة الماجستير ممن يحصلون على تقديرات "A".

ب. درجة الوسط الحسابي لصف الواحدة بعد الظهر في الامتحان النهائي ستكون أعلى بصورة دالة من درجة الوسط لصف السابعة مساءاً.

جـ. الدوام المسائي بدلاً من دوام بعد الظهر مفيد من حيث التقديرات بالنسبة للطالبات، وعائق بالنسبة للطلاب.

د. إذا أعيد نفس الاختبار القصير للإحصاء، الذي أعطي للطلبة في أول يوم دوام لهم، في آخر يوم للدوام، فسوف يتضح تقدم دال لدى الطلبة.

هـ. هناك علاقة بين النجاح أو الرسوب والتعبير عن الرضى أو عدمه بالنسبة للمقرر.

و. إن أية فروقات بين أوساط الدرجات، لصفوف دورة الخريف وصفوف دورة الصف لنفس المدرس، يمكن أن تعزى بسهولة إلى الصدفة.

ز. هناك أثر تفاعلي دال بين خبرة التدريس وخبرة البحث والنجاح في مقرر البحث.

21. يختلف الاختبار الاتجاهي عن الاختبار اللااتجاهي في أن الباحث في الاختبار الاتجاهي مهتم بالتغيرات التي تحدث في أي مما يلي:

أ. فقط في الاتجاه الإيجابي.

ب. فقط في الاتجاه السلبي.

جـ. في كلا الاتجاهين الإيجابي والسلبي.

د. في الاتجاه الإيجابي أو في الاتجاه السلبي وليس في كليهما.

22. استخدمت مجموعتان منتخبتان عشوائياً في تجربة، حصلت المجموعة 1 فيها على معالجة ولم تحصل المجموعة 2 عليها. وكانت فرضية الباحث هي أن وسط الأداء للمجموعة 1 سيكون أعلى من وسط الأداء للمجموعة 2. طبق الاختبار التائي على المعلومات أدناه ووضح ما إذا كان بالإمكان تأكيد فرضية الباحث.

	$\overline{X}$	N	Σx^2
المجموعة **1**	45.32	30	382.02
المجموعة **2**	41.78	30	264.32

23. يريد أحد الباحثين اختبار فرضية تقول إن معامل الارتباط بين المتغير أ والمتغير ب هو أكبر من صفر. وقد حصل على 21. = r بين المتغيرين باستخدام 22 فرداً. استخدم الجدول A.5 لإيجاد ما إذا كان بالإمكان رفض الفرضية على مستوى 05. (بطرفين).

24. البيانات أدناه هي لتصميم لتجربة (3×2). طبق التحليل متعدد العوامل للتباين وذلك لاختبار دلالة التأثيرات الرئيسية وأثر التفاعل.

الأعمدة

الصفوف	أ	ب	جـ
أ	25, 23, 20, 17, 15	22, 20, 18, 16, 14	20, 18, 16, 14, 12
ب	16, 14, 12, 10, 8	18, 16, 14, 12, 10	19, 18, 16, 14, 13

أرسم جدولاً تتضح فيه حواصل جمع المربعات (*SS*)، ودرجات الحرية (*df*) ومربعات الأوساط (*MS*) والقيم الفائية، ثم أجب عن الأسئلة التالية:

أ. أي من القيم الفائية دال وعلى أي مستوى؟

ب. كم فرضية صفرية اختبرت في هذه المسألة؟

جـ. كم من هذه الفرضيات يمكن رفضها؟

25. أوضح منطق الاختبار التائي.

26. لديك قائمة من طلبة الثانوية الذين أعطوا أرقاماً من 1 إلى 1000. استخدم جدول الأرقام العشوائية في الملحق لانتخاب عينة من 50 من القائمة الافتراضية. ادرج الأعداد المنتخبة للعينة.

الأجوبة ANSWERS

1- تعتبر العينة الأكبر حجماً، المستمدة عشوائياً، أكثر احتمالاً في تمثيلها للمجتمع الإحصائي من العينة العشوائية الأصغر. فالعينة الكبيرة التي يتم الحصول عليها بطريقة تسمح بتحيز منتظم لن تكون ممثلة أكثر من عينة متحيزة صغيرة.

2- للحصول على عينة طبقية تناسبية، قسم الـ 500 مدرس حسب نسب تمثيلهم في المجتمع الإحصائي كما يلي:

ابتدائية	3500 / 7000 × 500 = 50
متوسطة	2100 / 7000 × 500 = 150
ثانوية	1400/7000 × 500 = 100
مجموع العينة	500

3. رقم قائمة بجميع الطلبة ثم انتخب عينة عشوائية لعدد معين باستخدام جدول الأعداد العشوائية. أبدأ بنقطة عشوائية في الجدول ثم أصعد أو أنزل في الجدول وأدخل الطلبة الذين سحبت أرقامهم.

4. أ. جميع المشاركين في المحلة.
ب. المشاركون في خمس مدن.
جـ. (500) شخص جرت مقابلتهم.

5. عينات احتمالية: أ، جـ، د، ز.
عينات لااحتمالية: ب، هـ، و

6. أ. ليس هناك علاقة بين لون الشعر ودرجة الفكاهة لدى الشخص.
ب. يقم الباحثون في خطأ النوع 1 إن أوضحوا أن الشقراوات لديهم فكاهة أكثر من غير الشقراوات، أو إن الشقراوات لديهن فكاهة أقل من غير الشقراوات، عندما يكون في الواقع لكلا المجموعتين مقدار متساو من الفكاهة.
جـ. يقع الباحثون في خطأ النوع 2 إن فشلوا في الاستنتاج أن للشقراوات فكاهة أكثر أو أقل عندما يكون في الواقع لديهن ذلك.
د. الباحث مع مستوى دلالة 05..
هـ. الباحث مع مستوى دلالة 001..

7. ب

8. يجب أن تكون الفرضية الصفرية صحيحة ويجب أن يرفضها الباحث.

9. أ

10. د

11. ب

12. يوازن عواقب خطأ النوعين 1، 2.

13. خطأ النوع 1.

14. قد تكون النتائج بسهولة من عمل الصدفة. والدليل غير كاف لتسويغ نتيجة ما.

15. يتم في المعاينة الطبقية التمثيل في كل طبقة من خلال استخدام العشوائية، بينما في المعاينة الحصصية يتحقق التمثيل في الطبقات المختلفة بطريقة الحكم.

16. (1) حجم التأثير الذي يعتبر حداً فاصلاً بين الفرق ذي المعنى والذي بدون معنى، (2) مستوى الدلالة، (3) الاحتمال المطلوب لرفض فرضية صفرية خاطئة.

17. $N = (1/.10)^2\ (1.96 + 84)^2$ (لاحظ أنه مع حجم التأثير الصغير يلزم عدد كبير)

$= 100\ (7.84)$

$= 784$

18. 3.33 – غير دال.

19. 0.954 – غير دال.

20. أ. مربع كاي
ب. الاختبار التائي للأوساط المستقلة.
جـ. التحليل العاملي للتباين.
د. الاختبار التائي للأوساط غير المستقلة.
هـ. مربع كاي.
و. الاختبار التائي للأوساط المستقلة.
ز. التحليل العاملي للتباين.

21. د

22. $t = 4.11, df = 58$ دال على مستوى 001..

23. مع 20 درجة حرية و r يساوي (4227.) أو أكبر، يكون مطلوباً لرفض الفرضية الصفرية. وعليه يتم الاحتفاظ بالفرضية الصفرية.

24.

مصادر التباين	*SS*	*df*	*MS*	*F*
بين الأعمدة	0	2	0	0
بين الصفوف	120	1	120	11.34
تفاعل	80	2	40	3.78
بين المجموعات	200	5	40	3.78
داخل المجموعات	254	24	10.58	--

أ. بين الصفوف حسب01. وتفاعل حسب 05.
ب. ثلاثة
جـ. اثنان

25. تستخدم إحصاءات العينة لتحديد الفرق المتوقع بين وسطين الذي يحدث بالصدفة عندما تكون الفرضية الصفرية صحيحة. ويقارن الفرق الملاحظ مع الفرق المتوقع لتحديد احتمال أن يكون الفرق الملاحظ بسبب الصدفة وحدها.

26. ستختلف الإجابات.

REFERENCES المصادر

Ary, D., and Jacobs, L.C. (1976). Introduction to statistics. New York: Holt, Rinehart and Winston.

Borenstein, M., and Cohen, J. (1989). Statistical power analysis. Hillsdale, NJ: Lawrence Erlbaum.

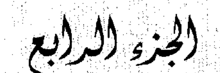
Cochran, W.C. (1985). Sampling Techniques (2d ed.). New York: Wiley.

Glass, G.V., and Hopkins, K.D. (1984). Statistical methods in education and psychology (2d ed.). Englewood Cliffs, NJ: Prentice Hall.

Jackson, P.W., and Lahaderne, H.M. (1967). Scholastic success and attitude toward school in a population of sixth graders. Journal of Educational Psychology, 58, 15-18.

Stuart, A. (1984). The ideas of sampling (3d ed.). New York: Macmillan.

Sudman, S. (1976). Applied sampling. New York: Academic Press.



أسس القياس

FUNDAMENTALS OF MEASUREMENT

* أدوات البحث.
* الصدق/ الصحة والثبات / الاستقرار



أدوات البحث

Tools of Research

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون الطالب قادرا على أن:

1. يوضح دور القياس في البحث.
2. يستخدم الكتب السنوية للقياسات العقلية للحصول على البيانات الضرورية في تقييم الاختبارات المقننة وغيرها من أدوات القياس.
3. يذكر الفرق بين الاختبار والمقياس.
4. يميز بين الاختبارات مرجعية المتوسط والاختبارات مرجعية الإتقان.
5. يميز بين مقاييس الاستعداد والتحصيل / الإنجاز.
6. يعرف تقييم الأداء ويناقش مزاياه ومساوئه.
7. يصف الخطوات المزمع اتباعها في أعداد مقياس (ليكرت) لقياس الاتجاهات.
8. يصف الخطوات المزمع اتباعها في إعداد مقياس ثيرستون.
9. يقارن خواص مقياسي ليكرت وثيرستون.
10. يصف خواص مقياس تمايز المعاني.
11. يوضح أنواع الأخطاء المألوفة في مقاييس التقدير.
12. يوضح مزايا ومساوئ مقاييس الشخصية حسب التقرير الذاتي.

13. يدرج خمسة إرشادات على الأقل، ينبغي على الباحث أن يتبعها عند استخدام الملاحظة المباشرة كأسلوب لجمع البيانات.
14. يعرف اختبار الموقف ويذكر متى يمكن استخدامه في البحث.
15. يوضح الخواص الأساسية للطريقة الإسقاطية ويسمى طريقتين إسقاطيتين مشهورتين على الأقل.

إن أحد أهداف البحث التربوي هو إحراز فهم أكبر للعلاقات بين المتغيرات في المجتمعات الإحصائية، مثلاً، قد يسأل المرء، ما هي العلاقة بين الذكاء والإبداع بين الأطفال بعمر السادسة؟ فلا يمكننا أنْ نلاحظ مباشرة الذكاء أو الإبداع، كما لا يمكننا أن نلاحظ مباشرة جميع أطفال السادسة. لكن ذلك لا يعني أننا يجب أن نظل نجهل هذه المسألة ومثيلاتها. فلدينا مؤشرات من شأنها أن تقرب المفاهيم البنائية للذكاء والإبداع، أي أن هناك أنواعا من السلوك الملاحظ نقبلها كمؤشرات صادقة للمفاهيم البنائية هذه. واستخدام المؤشرات لتقريب المفاهيم البنائية هو وجه القياس في البحث.

وتكون بعض القياسات مباشرة تماما وذلك باستخدام مؤشر واحد لتمثيل متغير معين. فمثلاً، بوسعنا قياس الخلفية التربوية الثقافية لشخص معين وذلك بسؤاله عن أعلى مستوى دراسي اجتازه. وبشكل مماثل، يمكن قياس متغيرات مثل مستوى الدراسة، القومية، الحالة الزوجية أو عدد الأطفال عن طريق مؤشر واحد، لأن مثل هذه المتغيرات تشير ببساطة إلى ظواهر واضحة جداً، ويؤمّن مؤشر واحد مقياسا مقبولا لها. أما المتغيرات الأخرى فهي، على أية حال، أكثر تعقيداً وأصعب قياساً. وفي هذه الحالات لا يعد استخدام مؤشر واحد مناسباً.

إن إحدى أهم المهام لدى الباحث هو انتخاب أو تطوير مقاييس وأدوات بوسعها أن تقيس سمات معقدة مثل الذكاء والإنجاز / التحصيل والشخصية والدافعية والاتجاهات والاستعداد والاهتمامات / الميول، ومفهوم الذات، وما إلى ذلك. وتستخدم أنماط معينة من هذه الأدوات لتكميم (التعبير بطريقة كمية) مثل هذه الصفات المختلفة.

وهناك طريقتان رئيسان للحصول على هذه القياسات: استخدم واحدة تم تطويرها، أو قم ببناء وسيلتك.

ولانتخاب أداة قياس، ينبغي على الباحث، أن ينظر إلى البحث الذي نشر حول مشكلته، كي يرى ما استخدمه الباحثون الآخرون لقياس المفهوم البنائي المعني. وسوف يوضح هؤلاء المؤلفون، بوجه عام، ما إذا كانت الأداة ناجحة أو ما إذا كانت إجراءات أخرى أفضل. أما المصادر الأخرى المفيدة لتحديد الأدوات المنشورة التي تخدم هدف الباحث، فهي (الكتب السنوية للقياسات العقلية والاختبارات المنشورة) التي تم وصفها في الفصل 3. وهناك مصدر آخر جيد للمعلومات حول الاختبارات، هو مجموعة اختبارات قسم الاختبار التربوي (ETS). فمجموعة اختبارات القسم هذا، عبارة عن مكتبة فيها أكثر من (16000) اختباراً بحثياً وتجارياً

وأدوات قياس أخرى مصممة لتأمين معلومات حديثة حول الاختبارات للباحثين التربويين. فهي تفهرس الاختبارات المنشورة وغير المنشورة. وهناك إضافة حديثة إلى هذه السلسلة، وهي الفهرس التراكمي للاختبارات بالمايكروفيش (75-1987) الذي يفهرس الاختبارات البحثية بتحديد المؤلف والعنوان والموضوع. ففي فهرس الموضوع هناك أكثر من (750) عنواناً تغطي كل شيء، من التفكير التجريدي إلى اتجاهات العمل.

وإذا تعذر على الباحثين إيجاد أداة متطورة سابقاً، فيجب عليهم عندئذ بناء أداتهم. وينطوي هذا الإجراء على تحديد واستخدام سلوك يمكن أن يعتبر مؤشراً على وجود المفهوم البنائي. ولغرض تحديد هذه المؤشرات، ينبغي على الباحثين أن يتحولوا أولاً إلى النظرية التي تقف وراء البحث. فالنظرية الجيدة سوف تشير؛ بوجه عام، إلى الكيفية التي سيتضح فيها المفهوم البنائي وكذلك التغييرات التي يمكن ملاحظاتها، أي أنها ستحدد طرق قياس المفهوم البنائي. فالباحثون يمكنهم أيضاً استخدام خبراتهم وإطلاعهم الواسع، كي يحددوا المؤشرات المناسبة للمفهوم البنائي. وفي الجزء التالي سوف نناقش باختصار بعض أنواع أدوات القياس المستخدمة في البحوث التربوية.

أدوات قياس MEASURING INSTRUMENTS

الاختبارات Tests

الاختبارات، أدوات قياس قيّمة للبحوث التربوية. فالاختبار جملة من المثيرات تقدم لشخص، بغية الحصول على استجابات تكون أساساً لتحديد درجات عددية. وهذه الدرجة، القائمة على عينة ممثلة لسلوك الشخص، هي مؤشر لمدى ما يحوزه المفحوص للخاصة التي يراد قياسها.

أما فائدة هذه الدرجات باعتبارها مؤشرات للمفهوم البنائي، موضع الاهتمام، فهي إلى درجة كبيرة دالة لموضوعية وصدق وثبات الاختبارات. وتشير الموضوعية إلى مدى الاتفاق بين الدرجات. فبعض الاختبارات، كاختبارات الاختيار من متعدد والصواب / الخطأ، توصف على أنها موضوعية لأن تسجيل الدرجات يتم بمقارنة إجابات الطالب مع مفتاح التصحيح، بمعزل عن أي قرار للمصحح. وتعتبر اختبارات المقالة أقل موضوعية لأن الدرجات تتأثر بأحكام وآراء المصححين. وسنناقش الصدق والثبات في الفصل 8.

اختبارات التحصيل Achievement Tests

في أغلب البحوث المعنية بفاعلية طرائق التدريس، يكون المتغير التابع هو التحصيل. وعليه، فاختبارات التحصيل مستخدمة بشكل واسع في البحوث التربوية، وفي الأنظمة المدرسية أيضاً. فهي تقيس مدى إتقان وكفاءة الأشخاص في مجالات مختلفة من المعرفة.

وتصنف اختبارات التحصيل بصورة عامة على أنها مقننة، وأخرى من صنع المدرس /الباحث. والاختبارات المقننة هي اختبارات منشورة ناتجة عن إعداد دقيق وماهر وتغطي أهدافاً أكاديمية واسعة مشتركة بين غالبية الأنظمة المدرسية. وهي اختبارات يتم من أجلها اشتقاق معايير/متوسطات(*) مقارنة، وتوطيد صدقها وثباتها، ووصف توجيهات لإجرائها وتصحيحها. وتوجد هذه التوجيهات في كتيبات يوفرها ناشرو الاختبارات.

بغية توطيد دعائم متوسطات هذه الاختبارات، فإن مبتكريها يجرون هذه الاختبارات على عينة ممثلة مناسبة تسمى المجموعة القياسية (Norm group). وقد يتم اختيارها لتمثل أمة برمتها أو ولاية أو مدينة أو مقاطعة أو مدرسة محلية. ويصبح الوسط الحسابي لمستوى دراسي معين في العينة المتوسط لذلك المستوى. أما المهارات المقاسة فهي ليست بالضرورة ما "كان يجب" تدريسه على أي مستوى، لكن استخدام المتوسطات يعطي التربويين أساساً لمقارنة مجموعاتهم مع تقدير الوسط الحسابي لجميع الأطفال في ذلك المستوى الدراسي. وتتوفر الاختبارات المقننة لموضوعات مدرسية خاصة مثل الرياضيات والكيمياء وكذلك بشكل حشيدات (بطاريات) شاملة لقياس عدة مجالات تحصيل. مثلاً يحتوي اختبار كاليفورنيا للتحصيل (CAT) على اختبارات في مجالات القراءة واللغة والحساب. وتشمل الاختبارات التتابعية للتقدم التربوي (STEP) على اختبارات في سبعة مجالات.

ولدى اختيار اختبار التحصيل، يجب على الباحثين أن يكونوا دقيقين في اختيار الاختبار بأن يكون موثوقاً ومناسباً، (يتمتع بالثبات والصدق) لقياس جانب التحصيل الذي يهتمون به. فيجب أن يتمتع الاختبار بالثبات والصدق بالنسبة لنوع الموضوعات المشمولة بالدراسة، فأحياناً لا يكون الباحث قادراً على انتقاء الاختبار إذ يتعين عليه استخدام ما اختاره النظام المدرسي. وتقدم (الكتب السنوية للقياس العقلي(**)) قائمة شاملة من مراجعات لاختبارات التحصيل المختلفة المتيسرة.

إذا تيسر اختبار لقياس السلوك المطلوب، وكان الثبات والصدق والقواعد كافية لهذه الغاية، عندئذ تكون هناك فوائد في استخدام أداة مقننة. فإضافة إلى توفير الوقت والجهد، سيدرك الباحثون فائدة استمرار إجراءات الاختبار – إذ يمكن مقارنة نتائج دراساتهم وتفسيرها إزاء نتائج دراسات أخرى تستخدم نفس الأداة. وعندما لا يعد استخدام اختبارات التحصيل المقننة مناسباً لأهداف معينة في دراسة بحثية معينة، فقد يبنى الباحثون اختباراتهم الخاصة بهم.

(*) يستخدم المصطلح "Norm" في القياسات التربية، غالباً، وهو يستخدم أحيانا بمعنى المتوسط (الوسط الحسابي أو الوسيط) ويترجمه بعضهم إلى "معيار". ويتم استخدام كلمة "معيار" في مواقف عدة بما يربك المتعلم. ونفضل استخدام "المتوسط" كمعيار في حالات الاختبارات المقننة إذ ييسر هذا الفهم للمتعلم وهو ما يشير إليه (Mehrens & Lehmann, 1980) في كتابهما (Standardized Tests in Education)، فالغرض من الاختبار هو تحديد موقع الفرد بالنسبة لأقرانه. (المراجع).

(**) يشير مصطلح "القياس العقلي" إلى قياس الذكاء: (المراجع).

فمن الأفضل بناء اختبار خاص بالباحث بدلاً من استخدام اختبار غير مناسب لأنه المتيسر، فحسب. وينبغي في هذه الحالة أن يهتم المرء بإعداد الاختبار خصوصاً فيما يتعلق بتحديد صدق وثبات الاختبار قبل استخدامه. وكمقترحات بشأن بناء الاختبار، قد يعود المرء إلى النصوص المتخصصة في القياس كتلك التي قام بتأليفها (Gronlund and Linee, 1990) و (Sax, 1989) و (Mehrens and Lehmann, 1991).

وحسب الأساس الذي قاد إلى نمط التفسير، قد تصنف الاختبارات المقننة التي يعدها المدرس بكونها مرجعية المتوسط، أو مرجعية الإتقان(*). فالاختبارات مرجعية المتوسط تسمح للمرء بمقارنة أداء الأفراد في الاختبار مع أداء الأشخاص الآخرين. ويفسر أداء الفرد في إطار موقعه النسبي في مجموعة محددة تعرف بمجموعة القياس / المقارنة. ويذكر الأداء في إطار المقاييس المئينية / المئوية والدرجات المعيارية وما أشبه .

ومن ناحية أخرى، تمكن الاختبارات مرجعية الإتقان، الباحث في وصف ما يمكن لشخص محدد القيام به دون الرجوع إلى أداء الآخرين. ويذكر الأداء في إطار مستوى الإتقان لمحتوى معين أو مجال مهارة.

ويتضح مستوى التفوق، بصورة نموذجية، بنسبة البنود التي كانت إجاباتها صحيحة. وتستخدم درجات فاصلة محددة مسبقاً لتفسير أداء الفرد. وقبل تصميم أدوات القياس يجب أن يحدد المرء نوع التفسير الذي يراد القيام به. ففي الاختبارات مرجعية المتوسط تنتخب البنود التي ستتمخض عن مدى واسع من الدرجات. فيجب على المرء أن يهتم بصعوبة البنود وقوة البنود بغية التمييز بين الأشخاص. أما في الاختبارات مرجعية الإتقان فتنتخب على أساس جدارتها، فحسب، لقياس مجموعة محددة من الأهداف التعليمية. وقد تكون هذه سهلة أو صعبة استناداً إلى ما يراد قياسه. وينصب الاهتمام الرئيسي على تحديد عينة ممثلة للبنود التي تقيس الأهداف الموضحة بحيث يمكن وصف الأداء الفردي مباشرة في إطار المعرفة والمهارات المحددة التي يمكن لهؤلاء الأشخاص تحصيلها.

قد تستخدم (معدلات النقط– GPA) لأفراد في مدارسهم كمؤشرات للنجاح والتحصيل الأكاديمي. وينبغي اتخاذ الحذر عند استخدام معدلات (النقط) لطلبة في أنظمة مدرسية مختلفة.

(*) هناك مصطلحان يشيران إلى هذين الصنفين "norm-referenced ،Criterion-referenced"، وقد شاعت ترجمة عربية تشير لهما "مرجعي المعيار، مرجعي المحك" وهاتان ليستا قريبتين من الفهم المباشر للمتعلم، ذلك أن كلمتي "معيار، محك" عامتان وتعنيان تقريبا المعنى ذاته إذ تشيران إلى "مجموعة العوامل التي يجري الحكم في ضوئها". لكن "معيار" في الأول يشير إلى السؤال: ما موقع الفرد بين أقرانه؟ ويفيد هذا بالبحث عن موقع الفرد بالنسبة للمتوسط، أما "محك" في الثاني فيشير إلى السؤال"كم أتقن؟ أي يسأل عن عدد عناصر الاختبار الأساسية التي أتقنها". ومن هنا، اقترحنا أن تكون الترجمة مباشرة "مرجعي المتوسط، مرجعي الإتقان". (المراجع)

ولا تعني الدرجات المقدرة بالحروف في المدارس المختلفة الشيء ذاته، بالضرورة، ولا يمكن الاعتماد عليها لإيضاح نفس درجة الإتقان والمهارة.

تقييمات الأداء Performance Assessments

إن معظم اختبارات التحصيل المألوفة هي اختبارات القلم والورقة التي تقيس الأهداف الإدراكية / المعرفية. كذلك، قد يرغب الباحث في بعض الحالات بقياس الأداء – أي ما يمكن أن يفعله الشخص وليس ما يعرفه. وأضحى تقييم الأداء بديلاً شائعاً للاختبارات التقليدية بين التربويين. فاختبار الأداء هو أسلوب يلاحظ فيه المرء ويقيم مباشرة أداء الفرد في واجب معين، و/ أو يحكم على الحاصل النهائي. ويتم الحكم على الأداء أو الحاصل في ضوء معايير محددة. ويحدد (Stiggins, 1992)، وهو أحد أوائل المؤيدين لتقييم الأداء، ذلك في إطار أربعة عناصر: 1) سبب التقييم. 2) أداء معين يراد تقييمه، 3) تمارين ناتجها الأداء، 4) إجراءات تقدير نظامية. فمثلاً، في استقصاء حول طريقة جديدة لتدريس العلوم، قد يرغب المرء في معرفة أثر الطريقة، ليس فقط على السلوك الإدراكي/المعرفي للطلبة، بل كذلك على تعلمهم إجراءات وأساليب المختبر المتنوعة أو قدرتهم على إكمال التجارب. في هذه الحالة يتطلب اختبار الباحث أداء الطلبة في واجب حقيقي أو استخدام معرفتهم ومهاراتهم في حل مسألة علمية. إن تقييم الأداء مهم في مجالات مثل الفن والموسيقى والاقتصاد المنزلي والخطابة العامة والعلوم التي تنطوي بصورة نموذجية على قدرة الفرد في القيام بشيء أو إنتاج شيء ما.

يشار أحياناً إلى تقييم الأداء بمصطلح (التقييم الصادق)، لكن (Wiggins, 1993) لا يعتبر جميع تقييمات الأداء صادقة. فلكي يكون التقييم جديراً بالمصداقية، يجب أن ينطوي على تقييم أداء الطلبة في أوضاع وواجبات حقيقة بدلاً عن أوضاع مصطنعة. ويوضح (Hart, 1994) أن التقييم يكون جديراً بالثقة، عندما ينطوي على نشاطات حقيقية ذات قيمة ومهمة وذات معنى. وينطوي التقييم الصادق على نشاطات متنوعة كالمقابلات الشفوية، والحقائب الكتابية، ومهام حل المسائل في مجموعات كما أوضحها Wiggins بقوله:

> **"إن التقييمات الصادقة تكرر تحديات ومعايير الأداء التي تواجه الكتّاب، أو رجال الأعمال، أو العلماء، أو قادة المجتمع أو المصممين أو المؤرخين. وهي تشمل كتابة مقالات، وتقارير، وإجراء بحوث جماعية أو فردية، وتصميم مقترحات ونماذج مطابقة، وجمع حقائب وما إلى ذلك (1989, p. 703).**

وننصح القارئ بالعودة إلى مقالة (Wiggins, 1993) للإطلاع على بحث جيد حول التقييم الصادق ومعايير الحكم على مصداقية اختبار معين.

بناء اختبار أداء Construction a Performance Test: عندما يتطلب تقييم أداء، ينبغي على الباحث اتباع ثلاث خطوات أساسية في بناء الاختبار:

1- إبدأ بعبارة واضحة للأهداف أو ما ينبغي على الأفراد القيام به في الاختبار، والشروط التي يخضع لها أداء المهمة. وينبغي على الباحث إيضاح ما إذا كانت هناك حدود زمنية وما إذا كانت المراجع متوفرة وما إلى ذلك.

2- أعط مسالة أو تمريناً لمنح الطلبة فرصة للأداء – إما محاكاة أو واجب فعلي. ينبغي أن تطلب من جميع الأفراد أداء نفس العمل.

3- قم بتطوير أداة (قائمة تدقيق، صيغة تقدير/ تصنيف أو ما يشابه) تسجل المعايير ذات الصلة والمراد استخدامها في تقييم الأداء و/أو النتاج. ولرفع الموضوعية إلى الحد الأعلى ينبغي أن تؤمن الأداة استخدام نفس المعايير لأداء أو نتاج كل فرد.

فوائد وقيود Advantages and Limitations: تؤمن اختبارات الأداء طريقة لقياس القدرات والمهارات التي لا يمكن إجراؤها بالاختبارات التحريرية. وهي، على أية حال، تتطلب كثافة في الوقت وعليه فهي باهظة التكاليف، سواء من حيث التطبيق أو وضع الدرجات.

اختبارات الاستعداد Aptitude Tests

تختلف اختبارات الاستعداد عن اختبارات التحصيل حيث تحاول الأولى قياس قدرة عامة بينما تحاول الثانية قياس مهارات في مجالات محددة. وتحاول اختبارات الاستعداد قياس قدرة الأفراد في فهم العلاقات وحل المشكلات وتطبيق المعرفة في سياقات مختلفة.

أشير سابقاً إلى اختبارات الاستعداد بكونها (اختبارات ذكاء) إلا أن المصطلح الأخير قد تلاشى في الاستخدام بسبب الجدل حول تعريف الذكاء وبسبب ميل الناس إلى ربط الذكاء بقدرة موروثة. فلا ينبغي اعتبار اختبارات الاستعداد كمقاييس للذكاء الفطري (أو المحْض). فالأداء في مثل هذه الاختبارات يعتمد إلى حد ما على خلفية الموضوع وتدريسه.

ويجد التربويون أن اختبارات الذكاء مفيدة وصحيحة بصورة عامة بهدف التنبؤ بالنجاح المدرسي. ففي الواقع يشار إلى العديد من الاختبارات على أنها (اختبارات استعداد دراسي) وهذا مصطلح يشير بصورة محددة إلى الوظيفة الرئيسية لهذه الاختبارات وهي التنبؤ بالأداء المدرسي. ومن اختبارات الاستعداد المشهورة هناك اختبار ACT (تقييم اختبار الكليات الأمريكية) واختبار SAT(*) (اختبار التقييم الدراسي) للمدارس الثانوية واختبار GRE (اختبار سجل الخريج) واختبار MAT (اختبار مقارنات ميلر) لطلبة التخرج في الكلية.

ويستخدم الباحثون كذلك اختبارات الاستعداد بشكل مكثف. فالاستعداد أو الذكاء هو

(*) يشير SAT هنا إلى (Scholastic Assessment Test) كما يشير أيضاً إلى (Scholastic Aptitude Test) أي اختبار استعداد دراسي. (المراجع)

في الغالب، متغير بحاجة إلى التحكم / الضبط في التجارب التربوية. ولضبط هذا المتغير يستخدم الباحث الدرجات من اختبار الاستعداد المدرسي. ومن بين العديد من الاختبارات المتوفرة، هناك اختبارات صممت للاستخدام مع أفراد وأخرى مع مجموعات.

اختبارات الاستعداد الفردية Individual Aptitude Tests: إن أوسع أدوات قياس الاستعداد على الصعيد الفردي هي اختبارات ستانفورد – بينيه واختبارات ويكسلر الثلاثة. فالاختبار الأول جاء نتيجة مراجعات عدة للأداة التي طورها في فرنسا عام 1905 الفرد بينيه(*) وتيودور سايمون اللذان أرادا تحديد الأطفال الذين لا يحتمل استفادتهم من التدريس الصفي الاعتيادي. واستخدم الاختبار أساساً لقياس العمر العقلي للفرد. ولاحقاً تم إدخال مفهوم معامل الذكاء (IQ). ويستمد هذا المعامل بتقسيم العمر العقلي (MA) على العمر الزمني (CA) وضرب النتيجة في مائة. إن المراجعة الحالية لاختبار ستانفورد – بينه تعطي أعماراً عقلية، لكنها لا تستخدم النسبة MA / CA لتحديد IQ. بدلاً عن ذلك، يتم الحصول على IQ بمقارنة أداء الفرد (درجته) مع المتوسطات التي يتم الحصول عليها من المجموعة العمرية للفرد من خلال استخدام درجات معيارية (أنظر الفصل 5). أما السمة الرئيسية لاختبار ستانفورد – بينيه فهي إعطاء قياس عام للاستعداد. إنها لا تحاول قياس قدرات منفصلة كما تفعل الاختبارات الأخرى. أما الاختبارات التي طورها ديفيد ويكسلر لقياس الاستعداد، فقد اتخذت أشكالاً عدة: مقياس ويكسلر لذكاء الأطفال – الطبعة الثالثة (WISC-III) ومقياس ويكسلر المنقح لذكاء الكبار (WAIS-R) ومقياس ذكاء ويكسلر المنقح لذكاء أطفال ما قبل المدرسة والابتدائية (WPPSI-R) الذي أُدخل للمجموعة العمرية البالغة 4 – 6.5 .

هناك خاصة مهمة لاختبارات ويكسلر إذ أنها مقسمة إلى اختبارات فرعية تمكن القائم بالاختبار من الحصول على درجتين لكل موضوع، واحدة تخص IQ الشفوي وأخرى لغير الشفوي. وهذه الاختبارات الفرعية تنقسم بدورها إلى مقاييس فرعية توضح أداء (من يجري اختبارهم) في واجبات محددة.

اختبارات الذكاء للمجموعات Group Tests of Aptitude: ينبغي أن يجري اختبارات ستانفورد – بينيه أو ويكسلر متخصص بالقياس النفسي على فرد معين، وهو إجراء باهظ الثمن من حيث المال والوقت. وهذه الاختبارات غير عملية، عندما تكون هناك حاجة إلى مقاييس استعداد لمجموعات كبيرة من الأفراد. وفي هذه الحالة يتم استخدام اختبارات جمعية. وقد تم

(*) كان الفرد بينيه (Alfred Binet) يهدف في اختباره، إلى تشخيص الأطفال الذين كان مستوى تحصيلهم الدراسي متدنياً وذلك لتقديم برامج علاجية ترفع من مستوى تحصيلهم. ولم يعتقد بينيه مطلقاً بأن اختباره يقيس سمة ثابتة أو فطرية للطفل. وعندما ترامى إلى سامعة بأن هناك من يرى أن الذكاء كم ثابت لا يمكن زيادته، اعترض بشدة قائلاً "ينبغي أن نحتج ونثور ضد هذا التشاؤم البهيمي" ويمكن لمن يؤد إثراء معلوماته حول هذا الموضوع العودة إلى كتاب (Rose, S. et. al: Not In Our Genes, 1984) – (المراجع).

تطوير الاختبار الجمعي الأول للقدرة العقلية خلال الحرب العالمية الأولى بهدف قياس قدرة الرجال في الخدمة العسكرية. وأحد أشكال هذا الاختبار ويسمى (اختبار ألفا للجيش) قد أطلق استخدامه مدنياً بعد الحرب، وأصبح النموذج لعدد من اختبارات المجموعات. وتتوفر اليوم العديد من اختبارات المجموعات للاستعداد العقلي. ومن بين الاختبارات الأكثر شيوعاً هو اختبار كاليفورنيا للنضج العقلي (CTMM) واختبار القدرات الإدراكية/ المعرفية (CAT) واختبار أوتس – لينون للقدرات المدرسية (OLSAT).

قياسات الشخصية Measures of Personality

إن الحصول على مقاييس الشخصية، مجال اهتمام آخر بالنسبة للباحثين التربويين. وهناك عدة أنواع مختلفة من مقاييس الشخصية يعكس كل واحد منها رأياً نظرياً مختلفاً. فبعضها يعكس نظريات سمة ونمط وأخرى تعود بجذورها إلى نظريات التحليل النفسي والدافعية. وينبغي أن يعرف الباحثون بدقة ما يريدون قياسه، ثم يقوموا بانتخاب الأداة ويولوا اهتماماً خاصاً بدليل صدقه. والنوعان الأكثر شيوعاً من مقاييس الشخصية في البحوث هي الاستبيانات والطرق الاسقاطية.

الاستبيانات Inventories

يجري في الاستبيان عرض مجموعة واسعة من العبارات على الأفراد تصف أنماطاً سلوكية ويطلب منهم إيضاح ما إذا كانت كل عبارة مميزة لسلوكهم، وذلك من خلال الإشارة بكلمة نعم، كلا، غير متأكد. ويتم حساب درجاتهم بإيجاد عدد الإجابات التي تتفق مع السمة التي يحاول من يطبق الاختبار قياسها. فمثلاً المصابون بالبارانويا(*) يتوقع لهم أن يجيبوا بكلمة نعم على عبارة "يتحدث الناس دائماً وراء ظهري" وبكلمة كلا على عبارة "أتوقع من الشرطة أن تكون عادلة ومعقولة". بالطبع، إن مثل هذه الإجابات على عبارتين فقط لن يوضح نزعات البارانويا. فمثل هذه الإجابات على نسبة كبيرة من العبارات يمكن اعتبارها، على أية حال، مؤشراً على البارانويا. وتقيس بعض الاستبيانات ذاتية التقرير، سمة واحدة فقط مثل "مقياس – F كاليفورنيا" الذي يقيس التسلط. وأخرى مثل استبيان الستة عشر عاملاً للشخصية، المعروف بمقياس (Cattle)، فإنه يقيس عدداً من السمات. وهناك استبيانات معروفة أخرى مثل استبيان مينوسوتا للشخصية متعددة الأطوار، ومسح غيلفورد – زمرمان للمزاج، وقائمة موني لتدقيق المشكلات، وجدول أدواردز للتفضيل الشخصي، ومؤشر مايرز – بريكز ومسح المصلحة القوية.

وتستخدم هذه الاستبيانات في البحوث التربوية للحصول على أوصاف السمات لمجموعات محددة، مثل المتخلفين دراسياً والمتسربين من المدارس وما إلى ذلك. وقد استخدمت في البحوث المعنية بالعلاقات المتداخلة بين سمات الشخصية ومتغيرات معينة مثل الاستعداد

(*) يشير مصطلح البارانويا (Paranoia) إلى "مرض العظمة أو الاضطهاد الكاذب" ويسمى "الهذاء" – (المراجع).

والتحصيل والاتجاهات. وللاستبيانات فوائد مثل الاقتصاد والبساطة والموضوعية. أما أغلب المساوئ فترتبط بمشكلة الصدق. فصدقها يعتمد جزئياً على مدى قدرة المستجيبين على قراءة وفهم البنود، وفهمهم لأنفسهم، وخصوصاً رغبتهم في إعطاء إجابات صريحة وصادقة. وكنتيجة لذلك، فإن المعلومات المستمدة من الاستبيانات قد تكون سطحية أو متحيزة. ويجب أن يؤخذ هذا الاحتمال بعين الاعتبار عند استخدام النتائج المستمدة من هذه الأدوات.

الطرق الإسقاطية Projective Techniques

تعتبر الطرق الإسقاطية مقاييس يطلب فيها من الفرد الاستجابة لمثير غامض، أو غير محدد المعالم. وهي تدعى إسقاطية لأنه يتوقع أن يسقط الشخص حاجاته ورغباته ومخاوفه وقلقه وما إلى ذلك. وفي ضوء تفسير واستجابات الشخص يحاول القائم بالاختبار بناء صورة شاملة لهيكل شخصية الفرد. ويستخدم الطرق الإسقاطية علماء النفس السريريين وذلك لدراسة وتشخيص الأفراد ذوي المشكلات الانفعالية. وهي لا تستخدم بشكل متكرر في البحوث التربوية وذلك بسبب ضرورة التدريب المتخصص لإدارة، وإجراء هذه الطرق، وتسجيل الدرجات، وكذلك بسبب الكلفة الباهظة للتطبيق على الفرد. وعلاوة على ذلك يشعر العديد من الباحثين أن صدقها لم يتأكد بشكل مقنع. وهناك طريقتان معروفتان من الأساليب الإسقاطية وهما TAT (اختبار تفهم الموضوع) وطريقة رورشاش. أما رورشاش فيستخدم بقع الحبر كمثير في حين يعرض على المستجيب في طريقة TAT صوراً ويطلب منه أن يروي قصة حول كل منها.

مقاييس الاتجاهات Attitude Scales

تستخدم مقاييس لقياس الاتجاهات والقيم والآراء وغيرها من السمات التي لا يتيسر قياسها بسهولة في اختبارات أو أدوات أخرى. فالمقياس (SCALE) هو مجموعة من الأصناف أو القيم العددية مخصصة للأفراد أو الأشياء أو لأنواع السلوك (Behaviors) بهدف قياس المتغيرات. وعملية تخصيص الدرجات لهذه الأمور، بغية الحصول على قياس للمفهوم البنائي، تدعى عملية القياس (Scaling). وتختلف هذه المقاييس (Scales) عن الاختبارات (tests) في كون أنها، خلافاً للاختبارات، لا تشير إلى النجاح أو الفشل، القوة أو الضعف. إنها تقيس مدى درجة السمة المعنية التي يعرضها الفرد. فمثلاً، قد يستخدم المرء مقياساً لقياس اتجاهات طلبة الكلية نحو دين معين، أو موضوع ما. ولقد تم تطوير عدد من أساليب القياس على مدى السنوات.

وتستخدم مقاييس الاتجاهات التي سنبدأ مناقشتها، استجابات على أسئلة في العادة – ثم تدمج هذه الاستجابات في درجة واحدة للمقياس. أما مقاييس التقدير، التي سنناقشها لاحقاً في هذا الفصل، فتستخدم الأحكام – التي يقدرها الفرد المعني، أو الملاحظ – وذلك لتحديد درجات للأفراد أو الأشياء الأخرى لقياس مفاهيم بنائية تنطوي عليها. وتعد اتجاهات الأفراد والجماعات ذات أهمية لدى الباحثين التربويين. وقد يُعرّف الاتجاه بالشعور / الأثر السلبي أو

الإيجابي نحو جماعة خاصة أو مؤسسة أو مفهوم أو أمر اجتماعي. ويفترض (قياس الاتجاه)، القدرة على وضع الأفراد على طول سلسلة (كم) متصلة للتفضيل – أو عدمه تجاه هذا الأمر. وينبغي على الباحثين، في الغالب، بناء مقاييسهم الذاتية لقياس الاتجاهات. وسنناقش في الصفحات التالية ثلاثة أنواع من مقاييس الاتجاهات: المقاييس الجمعية أو مقاييس ليكرت، مقاييس الفترات المتساوية ظاهرياً أو مقاييس ثيرستون، ومقاييس تمايز المعاني.

مقاييس ليكرت: طريقة التقديرات الجمعية

Likert Scales: Method of Summated Ratings

يعتبر مقياس ليكرت (1932) (أنظر الشكل 7.1) واحداً من أكثر الأساليب شيوعاً في قياس الاتجاهات. ويقيّم مقياس ليكرت الاتجاهات نحو موضوع معين عن طريق عرض جملة عبارات عن الموضوع، والطلب من المستجيبين إيضاح ما إذا كانوا يتفقون بشدة أو يتفقون فحسب أو ليس لديهم قرار، أو لا يتفقون أولا يتفقون أبدا. وتخصص لاستجابات موافق – غير موافق المتنوعة قيمة عددية ويتم الحصول على الدرجات الكلية للمقياس بجمع الاستجابات العددية المقدرة لكل بند، وتمثل الدرجة الإجمالية هذه اتجاه الفرد نحو الموضوع.

الشكل (7.1): مثال عن مقياس ليكرت

مقياس المسؤولية الاجتماعية (SRS)

1- ليس هناك فائدة من القلق بشأن الأحداث الحالية أو الشؤون العامة وليس بوسعي عمل شيء بشأنها على أية حال.

أوافق بشدة　　أوافق　　متردد　　لا أوافق　　لا أوافق بشدة

2- ينبغي على كل شخص إعطاء بعض من وقته لصالح مدينته أو بلده.

أوافق بشدة　　أوافق　　متردد　　لا أوافق　　لا أوافق بشدة

3- سيكون حال بلدنا أفضل لو لم تكن لدينا العديد من الانتخابات ولم يتعين على الناس الانتخاب في الغالب.

أوافق بشدة　　أوافق　　متردد　　لا أوافق　　لا أوافق بشدة

4- ليس خذلانك لأصدقائك أمراً بالغ السوء لأنه لا يسعك القيام بالشيء الجيد لكل شخص طوال الوقت.

أوافق بشدة　　أوافق　　متردد　　لا أوافق　　لا أوافق بشدة

5- من واجب كل شخص القيام بعمله على أفضل ما يكون.
أوافق بشدة أوافق متردد لا أوافق لا أوافق بشدة

6- سيكون حال الناس أفضل إذا ما استطاعوا العيش بعيداً عن الآخرين دون القيام بعمل أي شيء لهم.
أوافق بشدة أوافق متردد لا أوافق لا أوافق بشدة

7- أتطوع في المدرسة عادة في مشاريع معينة.
أوافق بشدة أوافق متردد لا أوافق لا أوافق بشدة

8- أشعر بالاستياء عندما أخفق في إنهاء عمل وعدت بإنهائه.
أوافق بشدة أوافق متردد لا أوافق لا أوافق بشدة

يجري بناء مقياس ليكرت عن طريق تجميع عدد كبير من الاستجابات العددية حول موضوع معين، يعبر نصفها تقريباً عن اتجاه مؤيد بشكل واضح ونصفها الآخر (غير مؤيد بشكل واضح). ولا تستخدم بنود محايدة في مقياس ليكرت. ومن المهم أن تكون هذه العبارات عينة ممثلة لجميع الآراء أو الاتجاهات المحتملة حول الموضوع. وقد يكون مفيداً، التفكير بكل الجوانب الفرعية (الأبعاد) للاتجاه، ثم كتابة بنود حول كل جانب منها. ولإيجاد مثل هذه المجموعة المتنوعة من البنود، قد يجد الباحث أن من المناسب الطلب من الناس المعروفين بإلمامهم بالموضوع المعين والاتجاهات المحددة حوله، كتابة عدد من العبارات السلبية والإيجابية. وتعد الكتابات الافتتاحية حول موضوع معين، مصادر جيدة كذلك، للعبارات المحتملة في مقياس الاتجاهات.

تقدم العبارات، إضافة إلى إضافة الاستجابات (وهي خمسة بشكل نموذجي) على شكل سلسلة كم متصلة للموافقة - عدم الموافقة، إلى مجموعة من الأفراد. ويجب سحب / اختيار هذه المجموعة من مجتمع إحصائي مشابه للمجتمع الذي سيستخدم المقياس عليه. وينبغي كذلك ترتيب العبارات عشوائياً وذلك لتجنب أي اتجاه عقلي معين في الاستجابة لدى الأفراد. ويتم توجيه الأفراد لاختيار صنف الاستجابة الذي يمثل رد الفعل لكل عبارة بأفضل ما يمكن: أوافق بشدة(SA)، أوافق (A)، متردد (U)، لا أوافق (D)، لا أوافق بشدة (SD).

ولغرض تدريج المقياس، يجب تحديد الأوزان النسبية لأصناف الاستجابات. فبالنسبة للبنود المفضلة أو التي تم التعبير عنها بشكل إيجابي، تخصص القيم العددية 5 ,4 ,3 ,2 ,1 على التعاقب لأصناف الاستجابات بدءاً بالنهاية المفضلة فمثلاً، الصنف (أوافق بشدة) مع عبارة مفضلة تحصل على وزن (5)، أوافق (4)، لا أوافق بشدة (1). أما بالنسبة للعبارات غير المفضلة أو التي

تم التعبير عنها بشكل سلبي فيتم عكس الوزن لأن عدم الموافقة مع عبارة غير مفضلة، مكافئ من ناحية سيكولوجية للموافقة على عبارة مفضلة، وهكذا فإن (أوافق بشدة) على العبارات غير المفضلة ستحصل على وزن (1) و (لا أوافق بشدة) على وزن (5).

(لا تظهر القيم الوزنية على مقياس الاتجاهات المقدم للمستجيبين). أدرس الأمثلة التالية المنتخبة من مقياس لقياس الاتجاهات نحو الرياضيات:

	SA	A	U	D	SD
الرياضيات موضوعي المفضل في المدرسة (+)	5	4	3	2	1
أنا أكره الرياضيات بأي شكل (−)	1	2	3	4	5

فالشخص ذو الاتجاه المفضل نحو الرياضيات يوافق على بند إيجابي ولا يوافق على بند سلبي، والشخص الذي يكره الرياضيات سيوافق على البند السلبي ولا يوافق على البند الإيجابي. فإذا أشار الشخص إلى (A) في أول بند أعلاه و (D) في التالي فإن مجموع الدرجات سيكون (8): (4+4) للبندين. ويدل مجموع أوزان كل البنود التي يتفحصها الشخص على إجمالي درجة الشخص. أما نظام الوزن، فيعني أن الدرجة العالية للمقياس (SA للوحدات المفضلة، SD للوحدات غير المفضلة)، توضح اتجاهاً إيجابياً نحو الموضوع. وأعلى درجة ممكنة هي 5N حيث (N) تمثل عدد البنود بينما تكون الدرجة الدنيا الممكنة هي 1N.

تحليل المفردات/ البنود Item Analysis: بعد إجراء مقياس الاتجاهات على مجموعة أولية من المستجيبين يجرى تحليل المفردة/ البند بغية تحديد أفضل البنود. وتحليل البند يعطي نموذجياً ثلاث إحصاءات لكل وحدة: (1) مؤشر تمييز البند، (2) عدد و/ أو نسبة المستجيبين لكل اختيار لكل وحدة، (3) وسط البند والانحراف المعياري. ويبين مؤشر تمييز البند المدى الذي يميز كل بند بين المستجيبين بنفس الطريقة التي تميزها الدرجة الإجمالية. فإذا كانت هناك درجات إجمالية عالية لمن أعطوا درجات عالية للبند، وكانت هناك درجات إجمالية متدنية لمن أعطوا درجات متدنية للبند، فإن البند، عندئذ يميز بالطريقة ذاتها الدرجات الإجمالية للمقياس. ويحسب مؤشر تمييز البند عن طريق ارتباط درجات البند بمجموع درجات المقياس وهو إجراء يتم عادة بالحاسوب. ولأجل الفائدة يجب أن يكون معامل ارتباط البند (0.25) على الأقل مع الدرجة الكلية . أما البنود التي لها معامل ارتباط متدن جداً أو لديها معامل ارتباط سلبي مع الدرجة الكلية، فيجب إقصاؤها لأنها تقيس الشيء ذاته كالمقياس الكلي، ومن ثم لن تساهم بقياس الاتجاه. بالتأكيد، سيرغب الباحث بتفحص البنود التي وجد أنها غير مميزة. وقد تكون البنود غامضة أو مزدوجة

التساؤل (أي تحتوي على معتقدين أو رأيين في عبارة واحدة). أو قد تكون عبارات واقعية لا تعبر بصورة حقيقية عن المشاعر بشأن الشيء. وتنقيح هذه البنود قد يجعلها قابلة للاستخدام. أما الإحصاءتان الآخريتان لتحليل البند (2، 3 أعلاه) فتشيران إلى المدى الذي يستخدمه المستجيبون للخيارات المختلفة. فالبنود التي ينتشر عليها المستجيبون بين أصناف الاستجابات مفضلة على البنود التي تتجمع عليها الاستجابات في صنف واحد أو صنفين.

وبعد اختيار البنود المناسبة، حسبما يشير إليه تحليل البنود، فينبغي على الباحث أن يجرب المقياس المنقح مع مجموعة مختلفة من الأفراد ثم يدقق الوحدات مرة أخرى لأجل التمييز.

الصدق Validity: يصعب في الغالب تحديد معايير تستخدم للتثبت من صدق مقياس الاتجاهات. لقد استخدم بعض الباحثين ملاحظات السلوك الفعلي، كمعيار للاتجاه المقاس، إلا أن هذا الإجراء لا يستخدم في الغالب بسبب صعوبة تحديد السلوك الذي يكون أفضل معيار للاتجاه وكذلك بسبب صعوبة ضمان قياسات السلوك.

إن إحدى أسهل الطرق، لاستخلاص دليل صدق، هو تحديد المدى الذي يتمكن فيه المقياس من التمييز بين مجموعتين يعرف عن أفرادها أن لهم اتجاهات مختلفة. وللتحقيق من صدق مقياس يقيس الاتجاهات نحو دين منظم، يعمل المرء على تحديد ما إذا كان المقياس قد ميز بين أفراد الكنيسة الناشطين والناس الذين لا يذهبون إلى الكنيسة أو ليس لديهم تبعية لأية كنيسة. كذلك، فإن صدق مقياس لقياس الاتجاهات نحو الإجهاض ينبغي أن يميز بين أفراد الجماعات المؤيدة للحياة وأفراد الجماعات المؤيدة للاختيار. و "تميز" تعني أن مجموعتين يتوقع منهما أن يكون الاختلاف بين وسطي الدرجات حسب المقياس، ذا دلالة. وكطريقة أخرى، لتحديد الصدق، تكون بإيجاد معامل الارتباط بين الدرجات على مقياس الاتجاهات، مع الدرجات المستحصلة على مقياس اتجاهات آخر يقيس المفهوم البنائي ذاته، ويكون متمتعاً بالصدق.

الثبات Reliability: يجب كذلك تحديد ثبات المقياس الجديد. والخطوة الأولى هي التأكد بأن للمقياس طول كاف - بمعنى أنه يشمل بنوداً تكفي لتأمين معاينة ممثلة لمجال الآراء كلها التي تدور حول موضوع الاتجاهات. ومع تساوي الأشياء الأخرى، فإن حجم معامل ارتباط الثبات يرتبط مباشرة مع طول المقياس. وعلى أية حال، يوضح (Mueller, 1986) أنه إذا كان بناء البنود جيداً فإن المقاييس التي لها قليل من البنود تتراوح ما بين 20-22 ستتمتع بثبات مقنع (في الغالب أعلى من 0.80). أما عدد البنود المطلوبة فيعتمد جزئياً على مدى وضوح موضوع الاتجاهات. فكلما كان الموضوع مجرداً كلما احتجنا لبنود أكثر.

وقد يود المرء أيضاً حساب مؤشر الثبات. وأفضل مؤشر لمقياس الاتجاهات هو معامل الارتباط ألفا (أنظر الفصل 8) الذي يقدم قياساً للمدى الذي ترتبط فيه جميع البنود إيجابياً مع بعضها، وتعمل معاً لقياس سمة أو خاصة (الاتجاه). وتحسب البرامج الحاسوبية، مثل SPSS

(الرزمة الإحصائية للعلوم الاجتماعية) و SAS (نظام التحليل الإحصائي)، بطريقة روتينية المعامل ألفا كمقياس للصدق.

مقاييس ثيرستون: طريقة الفترات التي تبدو ظاهريا متساوية
Thurstone Scales: Method of Equal-Appearing Intervals

طور ثيرستون طريقة لتحديد قيم قياس خاص لبنود الاتجاهات (Thurstorn & Chave, 1929). فبينما تقيّم مقاييس ليكرت الاتجاهات عن طريق سؤال المستجبين الإشارة إلى درجة موافقتهم – عدم موافقتهم حول كل من سلسلة من العبارات، فإن مقاييس ثيرستون تقيم الاتجاهات عن طريق عرض عبارات حول موضوع يتراوح من، مفضل جداً مروراً بالمحايد إلى غير مفضل جداً، ثم الطلب إلى المستجيبين باختيار العبارات التي تتطابق أكثر ما يكون تقريباً مع اتجاهاهم. وينطوي بناء مقياس ثيرستون على الخطوات التالية:

1. جمع عدد كبير من العبارات (50–100) التي تعبر بشكل واسع عن درجات التفضيل – وعدم التفضيل المختلفة نحو موضوع الاتجاه بما في ذلك العبارات المحايدة. وتعطى العبارات إلى عدد كبير من الخبراء (50 أو أكثر) ممن لديهم معرفة كافية بالموضوع كي يستطيعوا تصنيف العبارات في أحد عشر صنفاً على مدى بعد "مفضل – غير مفضل". فالصنف "أ" يحتوي على العبارات التي يحكم على كونها مفضلة جداً. ويحتوي الصنف "ب" على العبارات التالية الأقرب إلى للمفضلة جداً وهكذا. ويمثل الصنف السادس "و" الموقع الحيادي فيما يخص الاتجاه، بينما يحتوي الصنف الأخير "ك" على العبارات غير المفضلة جداً.

ك	ي	ط	ح	ز	و	هـ	د	ج	ب	أ
11	10	9	8	7	6	5	4	3	2	1
غير مفضل				محايــد			مفضل			

ولا ينبغي أن يرتبط ترتيب العبارات في أصناف باتجاهات الخبراء الذاتية نحو الموضوع النفسي بل ينبغي أن يعكس فقط تصورهم للعبارات المفضلة أو غير المفضلة.

2. بعد أن يقدر الخبراء جميع البنود بصورة مستقلة، يجرى إعداد لتوزيع تقديرات الخبراء لكل بند. وسوف يوضح هذا التوزيع عدد الخبراء الذين وضعوا بندا في كل من الأصناف الإحدى عشر. مثلاً، افترض أن عبارة عن الكنيسة قد وضعه في الصنف "أ" 4 خبراء وفي الصنف "ب" "28" خبيراً وفي "جـ" "32" وفي "د" "16" خبيراً. ويتم حساب قيمتين وهما الوسيط و Q (الانحراف الربيعي) لهذا التوزيع. أولاً أحسب الوسيط:

الصنف	قيمة الصنف	الخبراء
د	4	16
جـ	3	32
ب	2	28
أ	1	4
		80

الوسيط = 2.5 + (8/32) 1

= 0.25 + 2.5

= 2.75

(راجع المعادلة 5.1)

إن الوسيط للتقديرات هو (2.75). ويصبح هذا قيمة المقياس المخصصة لذلك البند. توضح قيمة المقياس وضع الوحدة على سلسلة الكم الموجبة – السالبة. ولغرض دراسة مدى الموافقة بين الخبراء نجري حساب مؤشر التشتت / التغير لكل بند. وقد استخدمنا مقياس التشتت Q، وهو الانحراف الربيعي الذي يساوي نصف الفرق بين المئيني الخامس والسبعين والمئيني الخامس والعشرين (أنظر الفصل 5). ويفضل Q على الانحراف المعياري لأنه لا يتأثر بالدرجات المتطرفة. والانحراف الربيعي للمثال أعلاه هو 0.65 = 2/(3.38-2.07). وسوف يؤدي الاتفاق العالي بين الخبراء، بشأن الكيفية التي تكون فيها العبارة مفضلة أو غير مفضلة، إلى قيمة متدنية لـ Q. أما الاتفاق الضعيف / المتدني بين الخبراء فسوف يؤدي إلى قيمة عالية لـ Q. فالبنود التي لها قيمة Q عالية ترفض في العادة لكونها غامضة لا يمكن إدخالها في المقياس.

3. بعد حساب قيم المقياس (الوسيط) و Q لكل عبارة فإن الخطوة التالية تكون باختيار العبارات التي تمثل نقاطاً على سلسلة الكم مفضل – غير مفضل الموزعة بالتساوي من 1-11 قدر الإمكان. وحسب المدى الذي تمثل فيه قيم المقياس زيادات متساوية فأنه يكون لدى المرء مقياس فترات تقريبي. وإن كان جوهرياً، لبندين أو أكثر القيمة ذاتها فإنه يتم اختيار البند الذي له قيمة Q الأصغر. ويتم وضع البنود في سياق عشوائي بشكل نهائي، وقطعاً، لا يشار إلى القيم على الشكل نفسه.

يبين الجدول 7.1 عشر وحدات أخذت من مقياس ثيرستون لقياس اتجاهات نحو الكنيسة. وكانت قيم المقياس ذات هدف توضيحي فحسب، ولا تظهر في الشكل النهائي.

ولدى تطبيق مقياس ثيرستون، يوصي الباحث المستجيبين بالتدقيق في العبارات التي يوافقون عليها فقط. ودرجة اتجاه الفرد هو متوسط قيمة المقياس (الوسط أو الوسيط) للعبارات

المدققة. وتضع، درجة المتوسط، الفرد على سلسلة الكم مفضل - غير مفضل فيما يخص موضوع الاتجاه، فإذا مثلاً، وافق المستجيب على عبارات لها قيم مقياس 1.5، 2.3، 3,3، 4.5 في الجدول 7.1 فإن درجة اتجاهه ستكون 2.9 (الوسط) التي توضح اتجاها / موقفاً مفضلاً نحو الكنيسة

الجدول 7.1 مقياس الاتجاهات نحو الكنيسة

قيمة المقياس	العبارة
1.5	اعتقد أن العضوية في الكنيسة شيء مهم للعيش بأفضل ما يكون.
2.3	أجد خدمات الكنيسة مريحة وملهمة.
3.3	أتمتع بالكنيسة لأن روح الصداقة موجودة هناك.
4.5	أؤمن بما تعلمه الكنيسة، ولكن مع تحفظ عقلي.
5.6	أشعر أحيانا أن الكنيسة والدين ضروريان وأحياناً أشك بذلك.
6.7	أؤمن بالإخلاص والخير دون أية مراسيم كنسية.
7.4	أعتقد أن الكنيسة تفقد مكانتها مع تقدم التربية.
8.3	أعتقد أن تدريس الكنيسة مصطنع بحيث ليس له أهمية اجتماعية كبيرة.
9.6	اعتقد أن الكنيسة عائقاً عائق للدين لأنها لا تزال تعتمد على السحر والخرافة والأسطورة.
11.0	اعتقد أن الكنيسة طفيلية على المجتمع.

قد يعتبر مقدار الانتشار في قيم بنود الاتجاهات للمقياس التي يدققها المستجيب كمقياس للمدى الذي يعرّف فيه المستجيب الاتجاه بشكل واضح. فالشخص ذو الموقف المعرف جيداً نحو موضوع معين يتوقع له أن يدقق البنود القريبة جداً من قيمة المقياس فحسب. ولو انتشرت استجابات الشخص بشكل واسع على وحدات غير متجاورة، فمن المحتمل أن يكون لدى المستجيب اتجاها غامضاً أو سيئ التحديد.

مقارنة مقاييس ليكرت وثيرستون

Comparison of Likert And Thurstone Scales

هناك فوائد لكل من مقاييس ليكرت وثيرستون. فمقاييس ليكرت هي أسهل بناءاً كما وجدت عدة دراسات أنها أكثر ثباتاً حتى مع بنود أقل.

ولأن لمقياس ثيرستون نقطة حيادية حقيقية، فإن الفائدة الرئيسية له هي أن يوسع المرء أن

يقدم تفسيراً مطلقاً لدرجات المقياس، بينما يسمح مقياس ليكرت بتفسير الدرجات على أساس نسبي فحسب. ويسمح مقياس ليكرت بترتيب الأشخاص حسب تفضيلهم لاتجاه نحو موضوع معين إلا أنه لا يتيح سوى معلومات محدودة لمقدار مدى التفضيل بين شخص وآخر.

إن أكبر عيوب مقياس ثيرستون هو مقدار العمل الذي ينطوي عليه بناء المقياس. فيجب أن يكون لدى المرء مجموعة خبراء يقدرون التفضيل النسبي للعبارات كما يجب التأكد بأنهم يقدرون العبارات بشكل مستقل عن اتجاهاتهم نحو الموضوع. ويمكن أن يكون ذلك عملاً صعباً.

وتتوفر العديد من مقاييس ليكرت وثيرستون حول العديد من الموضوعات. وينبغي على الباحثين أن يبحثوا أولاً عن مقياس منشور قبل محاولتهم بناء مقياسهم الذاتي.

مقاييس تمايز المعاني / التباين المعنوي Semantic Differential Scales

هناك طريقة أخرى لقياس الاتجاهات وهي أسلوب تمايز المعاني الذي قام بتطويره (Osgood, Suci, and Tannenbaum, 1967).

يستند تمايز المعاني على افتراض أن للأشياء نوعين مختلفين من المعنى لدى الأشخاص، وهما المعنى الدلالي المعجمي الذي يشير إليه المعجم والمعنى الدلالي التضميني الذي يشير إلى الترابط أو الإيحاءات التي تستدعيها الكلمة. ويمكن للمرء أن يوضح المعنى المعجمي لشيء ما بسهولة أكثر من المعنى التضميني. ومن الممكن، على أية حال، قياس المعنى التضميني لشيء معين بصورة غير مباشرة، عن طريق الطلب إلى الأشخاص بتقدير الشيء عن طريق استخدام عدد من الصفات الثنائية القطب. وهكذا فإن معنى شيء لدى شخص سيكون نمط تقديراته لذلك الشيء على مقاييس الصفات الثنائية القطب.

لقد وجد أوسكود وزميليه من خلال التحليل العاملي (أنظر الفصل 11) ثلاث مجموعات (عناقيد) من الصفات: تقييمية (المكونة من صفات مثل جيد وسيئ أو قيم وتافه)، واقتدارية (المكونة من صفات مثل قوي وضعيف أو ثقيل وخفيف) ونشاطية (المكونة من صفات مثل نشيط وسلبي أو سريع وبطيء). ومن بين المجموعات الثلاث يظهر البعد التقييمي في كونه الأكثر أهمية في قياس الاتجاهات.

ويتم بناء مقاييس تمايز المعاني عن طريق اختيار أزواج من الصفات التي تمثل البعد التقييمي. وتقدم أزواج الصفات هذه على طول سبعة أصناف استجابة للمقياس، ويوجه المستجيب لوضع × في أحد الفراغات السبعة لإيضاح المدى الذي تصف عنده الصفة الشيء. فمثلاً، افترض أن المرء يرغب في قياس اتجاهات طلبة مدرسة ثانوية نحو المدرسة:

المدرســـــة								
سيئ	____	____	____	____	____	____	____	جيد
نشيط	____	____	____	____	____	____	____	سلبي
ذكي	____	____	____	____	____	____	____	غبي
سار	____	____	____	____	____	____	____	كئيب
تافه	____	____	____	____	____	____	____	قيم
صلب	____	____	____	____	____	____	____	طري
ثقيل	____	____	____	____	____	____	____	خفيف
ضعيف	____	____	____	____	____	____	____	قوي
سريع	____	____	____	____	____	____	____	بطيء

لاحظ في المقياس أن أزواج الصفات يتم إدراجها في كلا الاتجاهين بغية تقليص "استجابات اتجاهية" أي الميل لتفضيل بعض المواقع في قائمة الخيارات. فقد يكون لدى الفرد ميل لاختيار النهاية اليمنى المتطرفة فيشير إلى ذلك الموقع لكل بند. وعلى أية حال إذا ما تغير اتجاه المقياس بطريقة عشوائية بحيث لا تكون النهاية اليمنى الاستجابة الأكثر تفضيلاً فإن الشخص فيتم إجباره على قراءة كل بند والاستجابة حسب محتواها وليس حسب تفضيل موقعها. وفي إطار تسجيل درجات مقياس تمايز المعاني فإن الدرجات تحدد من 1-7 حيث تمثل 7 الاستجابة الأكثر إيجابية. وهكذا، في الوحدة الأولى أعلاه سوف يحصل الموقع الأول (سيئ) على الدرجة (1) والموقع الثاني على (2) ويحصل (جيد) على 7. وفي الوحدة الثانية تنعكس الدرجات، ف (نشيط) يحصل على 7 درجات و (سلبي) على درجة واحدة. أما التقدير لجميع البنود فيكون بجمع الدرجات وذكر متوسطها.

مقاييس التقدير Rating Scales

إن إحدى أدوات القياس الأوسع استخداماً هو مقياس التقدير. وتشمل مقاييس التقدير على تقييم شخص معين لسلوك أو أداء شخص آخر. ونموذجياً، يطلب من المقدّر أن يضع الشخص الذي يجرى تقييمه على درجة معينة في سلسلة كم متصلة أو في صنف يصف سلوكه المميز. وتلحق القيمة العددية بالدرجة أو الصنف. ويفترض أن المقدّرين على علم بالسلوك النموذجي للفرد. وتستخدم التقديرات في البحوث حول نمو الأطفال وحول العديد من أوجه السلوك.

الشكل (7.2): مثال على مقياس بياني

	متدني	وسط	عالي
مظهر شخصي			
قبول اجتماعي			
مهارات الكلام			

وثمة أنماط مختلفة من مقاييس التقدير. أما المقياس الأوسع استخداماً فهو المقياس البياني، الذي يوضح فيه المقدر قيمة التقدير بوضع إشارة في نقطة مناسبة على خط أفقي يمتد من نقطة سلوك متطرفة في السؤال إلى الأخرى. ويعد الشكل 7.2 مثالاً نموذجياً للمقياس البياني. ويمكن للمقدر أن يشير إلى أية نقطة على الخط المتصل. وتعين بعض المقاييس قيماً عددية للنقاط الوصفية. وتدعى هذه "مقاييس التقدير العددي". فيمكن لبند المهارات الكلامية في الشكل 7.2أن يبدو في المقياس العددي:

7	6	5	4	3	2	1
أحد أفضل المتكلمين			متكلم اعتيادي			أحد أسوأ المتكلمين

وهناك نوع ثان لمقياس التقدير وهو المقياس التصنيفي الذي يتكون من عدد من الأصناف المرتبة في سلسة منتظمة. ويستخدم خمسة إلى سبعة أصناف في الغالب. ويختار المقدّر الصنف الذي يكون أفضل ما يميز سلوك الشخص الذي يجرى تقييمه. افترض أنه يراد تقدير قدرات تلميذ، وإن إحدى الصفات المراد تقديرها هي الإبداع. وفيما يلي واحد من بنود التصنيف:

ما مدى إبداع الشخص؟ (حدد واحدة)

مبدع بشكل استثنائي ________
مبدع جداً ________
مبدع ________
غير مبدع ________
غير مبدع أبداً ________

وبغية تقديم معنى أكبر، تستخدم أحياناً عبارات وصفية موجزة لتكوين الأصناف في هذا النوع من المقاييس. وتسهم الأصناف المعرفة بشكل واضح، في ثبات التقديرات. فمثلاً:

ما مدى إبداع الشخص؟ (حدد واحدة)

له أفكار مبدعة دائماً ________

له العديد من الأفكار المبدعة ________

له أفكار مبدعة أحياناً ________

نادراً ما تكون له أفكار مبدعة ________

وباستخدام المقاييس البيانية والتصنيفية، يصدر المقدرون أحكامهم دون مقارنة مباشرة للشخص الذي يجرى تقديره مع الأشخاص أو المجموعات الأخرى.

وفي مقاييس التقدير المقارن، يوصى المقدرون بإعطاء أحكامهم من خلال إسنادهم المباشر إلى مواقع الآخرين الذين قد يقارن الشخص بهم. وتحدد المواقع على مقياس التقدير في إطار مجتمع إحصائي معين ذي سمات معروفة. ويتم إيضاح مقياس التقدير المقارن في الشكل 7.3. فمثل هذا المقياس قد يستخدم في انتخاب المتقدمين للقبول في كلية الدراسات العليا. ويطلب من المقدّرين الحكم على قدرة المتقدم في الدراسات العليا، بمقارنته مع جميع الطلبة الذين يعرفهم المقدر. ولكي يكون التقدير صادقاً لابد أن يكون لدى الحَكَم إدراك لمدى وتوزيع قدرات المجموعة الكاملة من طلبة الدراسات العليا.

الشكل (7.3): مثال عن مقياس تصنيف مقارن

مجال الكفاية / الامتداد (المراد تقديرها)	منخفض عادة	أسوأ من اغلب الطلبة	متوسط تقريبا بين الطلبة	أفضل من أغلب الطلبة	الأفضل حقاً	غير قادر على الحكم
1. هل يقدم هذا الشخص دليلاً على أهداف تأهيل واضحة المعالم وجديرة بالتقدير.						
2. هل يعالج هذا الشخص مشكلات بطريقة بناءة						
3. هل يتخذ النقد حسن النية ويستخدمه بشكل بناء						

أخطاء في التقدير Errors in Rating

إن جميع أساليب التقدير عرضة لخطأ كبير يقلل من صدقها وثباتها. ومن بين أكثر الأخطاء المنتظمة شيوعاً في تقدير الناس هو (تأثير الهالة) الذي يحدث عندما يسمح المقدرون للانطباع المعمم عن الفرد من التأثير على التقدير الذي يجري على جوانب محددة من السلوك. وينتقل هذا الانطباع العام من بند في المقياس إلى آخر. فمثلاً، قد يقدر معلم تلميذاً ذا عمل أكاديمي جيد بأنه متفوق في الذكاء والشعبية والإخلاص والمثابرة وجميع جوانب الشخصية. أو إن كان للمرء انطباع غير مُرض بشكل عام عن شخص معين فإن من المحتمل له أن يقدر ذلك الشخص بكونه متدنياً في جميع الأوجه.

هناك نوع آخر من الخطأ يدعى (خطأ الكرم) الذي يشير إلى الميل في تفسير الشك لصالح الفرد، أي عندما لا يكون المقدرون متأكدين، فمن المتوقع لهم أن يتحدثوا بشكل مُرض عن الفرد الذي يقومون بتقديره.

وبعكس ذلك، هناك (خطأ الصرامة) الذي يعد ميلاً نحو تقدير كل الأفراد بشكل متدن جداً في جميع الصفات. وهناك مصدر خطأ آخر هو (خطأ النزعة المركزية) الذي يشير إلى النزوع نحو تجنب أي تطرف وتقدير كل الأفراد في وسط المقياس.

إن إحدى طرق تقليص مثل هذه الأخطاء تكون في تدريب المقدرين بشكل عميق قبل الطلب منهم إجراء التقديرات. وينبغي إبلاغهم حول احتمال حدوث هذه الأنواع من الأخطاء.

ومن المهم أن يكون لدى المقدرين وقت كاف لرصد الفرد وسلوكه قبل إجراء التقدير. وكطريقة أخرى لتقليص الخطأ تكون بالتأكد من أن السلوك المراد تقديره، ونقاط مقياس التقدير، معرفة بشكل واضح. وينبغي وصف نقاط المقياس في إطار أنواع السلوك الواضح الذي يمكن ملاحظته، وليس في إطار أنواع السلوك الذي يتطلب الاستنتاج من جانب المقدّر.

ويزداد ثبات التقديرات عادة عن طريق العديد من المقدرين الذين يجرون تقديراتهم للفرد على نحو مستقل. ويتم تجميع التقديرات المستقلة أو إيجاد متوسطها للحصول على الدرجة النهائية.

أساليب القياس الاجتماعي Sociometric Techniques

تستخدم هذه الأساليب لدراسة تفاعلات الأشخاص ضمن المجموعات الاجتماعية. فالأجراء الأساسي، رغم إمكانية تعديله في عدة طرق، ينطوي على الطلب إلى أفراد المجموعة إيضاح خياراتهم الأولى والثانية والتالية لرفاقهم على أساس معيار معين وذلك عادة من أجل نشاط معين. فمثلاً، قد يطلب من أي طفل في مجموعة القراءة اختيار طفلين آخرين يود أن يدرس

معهما، أو يجلس بقربهما، أو يتناول الغداء معهما، أو يلعب معهما بعد المدرسة. إن طريقة القياس الاجتماعي تنصب أساساً على دراسة الخيارات التي يقوم بها كل شخص في المجموعة. ويتم تعيين مواقع الخيارات المستحصلة على ما يدعى "المبيان الاجتماعي" أو (Sociogram) الذي يبين نمط العلاقات بين الأشخاص في مجموعة معينة. وفي المبيان الاجتماعي الموضح في الشكل 7.4 قد يشار إلى سامي وهو الشخص الأكثر اختياراً على أنه (النجم). لاحظ أن كريم وهبة وهدى اختاروا بعضهم. وهذا يمثل زمرة، أي ثلاث أشخاص أو أكثر ممن يختارون بعضهم بالتبادل، بينما لم يحصل أوس على أية خيارات فهو منعزل. إن الخيارات التي كشف عنها في المقياس الاجتماعي يمكن تحويلها إلى "كم" واستخدامها لأهداف بحثية.

وتستخدم طرق القياس الاجتماعي بشكل واسع في بحوث علم النفس الاجتماعي وفي البحوث التربوية حيث قد تدرس حالات القياس الاجتماعي حسب علاقتها بالمتغيرات الأخرى كالقدرة العقلية والإنجاز / التحصيل وتفضيل المدرسين للأطفال.

الشكل 7.4 المبيان الاجتماعي لمجموعة

كريم
هبة
هدى
فرح
جميل
سامي
ياسمين
أوس

الرصد المباشر Direct Observation

في العديد من الحالات يعد الرصد المباشر المنتظم للسلوك طريقة القياس المفضلة. أما هدف الرصد المباشر فهو تحديد المدى الذي يكون فيه سلوك معين جلياً، فيحدد باحث السلوك المعني ويبتكر إجراءً منتظماً لتحديد وتصنيف وتسجيل السلوك في وضع طبيعي أو مخطط له. ويستخدم الرصد المنتظم بشكل واسع في البحوث على الأطفال الصغار وعلى أطفال ما قبل المدرسة. كما

يستخدم الرصد في البحوث النوعية والكمية. وعندما يتم الرصد في محاولة للحصول على صورة شاملة لوضع معين، ويكون حاصل هذا الرصد مجموعة ملحوظات أو سرد قصص، يكون البحث نوعياً. (أنظر الفصل 13). وفي إطار الفصل الحالي، يستخدم الرصد كمصدر كمي للبيانات - لأن نواتج استخدام أدوات الرصد المتنوعة تكون أعداداً.

تعد دراسة (Urban, 1943) الكلاسيكية بعنوان "التغيرات السلوكية الناجمة عند دراسة الأمراض السارية" مثالاً كلاسيكياً لاستخدام الرصد المباشر في وضع طبيعي. فقد سجل الراصدون عدد حالات السلوك غير المفضل مثل وضع الأصابع أو الأشياء الأخرى في الفم، وكذلك عدد أنواع السلوك المفضلة مثل استخدام المرء منديله عند السعال أو العطاس. وباتباع ذلك، انتخبت مجموعة تجريبية وتم تدريسها في دورة، لمدة ستة أسابيع حول الأمراض السارية، صممت لتغيير سلوكهم الواضح ولتزويدهم بالمعلومات الحقيقية والفهم. وفي نهاية الدورة تم تسجيل أنواع السلوك مرة أخرى. وقد وجد أن أنواع السلوك غير المفضل قد تقلص بشكل كبير حيث ازداد أنواع السلوك المفضل بشكل كبير في المجموعة التجريبية، بينما كان هناك تغير ضئيل في المجموعة الضابطة التي لم تأخذ الدورة. وقد وجد الرصد لمدة (12) أسبوعاً أن هذه الفروقات بين المجموعات التجريبية والضابطة، بقيت قائمة.

هناك خمسة خطوات أولية مهمة ينبغي اتخاذها في إعداد الرصد الكمي المباشر:

1- **يجب اختيار مظهر السلوك المرصود.** ونظراً لأنه لا يمكن جمع البيانات حول كل شيء يحدث فأنه يجب على الباحث أن يقرر مسبقاً أنواع السلوك التي يسجلها والأنواع التي لا يسجلها.

2- **يجب تعريف أنواع السلوك التي تقع ضمن صنف مختار بشكل واضح.** فيجب أن يفهم الراصدون الأفعال التي سيصنفوها كالسلوك التعاوني مثلاً أو السلوك الأناني.

3- **يجب تطوير نظام كلي للرصد / الملاحظة.** يجب على الباحث أن يقرر طريقة مقننة لتعداد أنواع السلوك الملاحظ. مثلاً يجب التحديد مسبقاً ما إذا كان يراد عد فعل ورد الفعل عليه باعتبار ذلك حادثاً واحداً للسلوك المرصود أو حادثين. أما الطريقة المقترحة فهي تقسيم فترة الرصد إلى أجزاء زمنية مختصرة ثم تسجيل في كل فترة - لنقل عشرة ثوان - فيما إذا كان الشخص أبدى السلوك أو لم يبده.

4- **يجب تطوير إجراءات محددة لتسجيل السلوك.** إن ذاكرة أغلب الراصدين لا يُعتمد عليها بما يكفي لإجراء بحث مفيد. وأفضل حل هو نظام تشفير يسمح بتسجيل مباشر لما يرصد وذلك باستخدام حرف أو رقم واحد بدلاً من نظام سردي يأخذ الكثير من وقت وانتباه الراصد.

5- **يجب تدريب من سيقومون بالرصد/ الملاحظة.** فالتدريب وفرصة الممارسة ضرورية كي يستطيع الباحث أن يعتمد على الراصدين لمتابعة إجراء محدد في رصد وتفسير وذكر الملاحظات. ويعتبر متابعة/دراسة الراصدين لشريط فيديو، ثم مناقشة النتائج أسلوب تدريب جيد.

الشكل (7.5): أصناف فلاندرز لتحليل التفاعل (FIAC)

حديث المعلم

1. **يقبل المشاعر:** يقبل ويوضح اتجاه أو شعور تلميذ بطريقة غير تهديدية. وقد تكون المشاعر إيجابية أو سلبية. ويشمل ذلك التنبؤ بالمشاعر واستدعائها.
2. **يمتدح ويشجع:** يمتدح ويشجع عمل أو سلوك التلميذ. يلقي النكات التي ترخي التوتر ولكن ليس على حساب شخص آخر، وهز الرأس أو قول "نعم، آه" أو "تابع" هي أمور مشمولة.

استجابة:

3. **يقبل أفكار التلاميذ أو يستخدمها:** يوضح، يبني أو يطور الأفكار المقترحة من تلميذ. ويشمل ذلك توسيعات المدرس لأفكار التلميذ، ولكن، بينما يعرض المعلم الأكثر من أفكاره، فإنه يسعى للانتقال إلى الصنف (5).

---

4. **يطرح أسئلة:** يطرح سؤالاً حول محتوى أو إجراء يستند إلى أفكار المعلم بهدف أن التلميذ سيجيب عليه.

---

5. **يحاضر:** يقدم حقائق أو آراء حول محتوى أو إجراءات، ويعبر عن أفكاره، ويعرض توضيحه أو يستشهد بمرجعية عبر التلميذ.
6. **يعطي توجيهات:** يعطي توجيهات أو أوامر يتوقع أن يتماشى معها التلميذ.

مبادرة :

7. **ينتقد أو يسوّغ السلطة:** يطرح عبارات يراد منها تغيير سلوك التلميذ من نمط غير مقبول إلى نمط مقبول، ويوبخ فرداً بسبب قيامه بعمل ما، موقف متطرف يتخذه.

حديث التلميذ:

استجابة

8. **حديث التلميذ – واستجابته:** حديث التلاميذ استجابة للمعلم. يبادر المعلم بالاتصال، ويطري عبارة التلميذ أو يبني وضعاً. حرية التعبير عن الأفكار محددة.

مبادرة

9. **حديث التلميذ – ومبادرته:** حديث التلاميذ الذي يبادرون به. أنهم يعبرون عن أفكارهم ويبادرون بموضوع جديد، ولديهم الحرية في تطوير آرائهم وخط أفكارهم، كطرح أسئلة متعمقة بما يتجاوز ما هو قائم.

الصمت

10. **صمت أم ارتباك:** وهذه فترات صمت قصيرة وفترات ارتباك التي لا يمكن فيها للراصد فهم ما يجري.

وهناك تطبيق آخر للرصد المباشر يقع ضمن دراسة السلوك الصفي، وقد طور (Flanders, 1965) نظام تشفير واسع الاستخدام لتسجيل الرصد الصفي. ويؤمن نظام فلاندرز الموضح في الشكل 7.5 أصنافاً شاملة ولا يحوي أي اثنين منها عنصراً مشتركاً، ويمكن تسجيل كل واحدة منها برقم واحد. والراصدون المدربون قادرون على تسجيل رقم كل ثلاثة ثوان. وبالوسع تحليل سلسلة الأرقام الناتجة بسهولة حيث لا يكون هنا سجل بنسب السلوك اللفظي في كل صنف، فحسب، بل نتزود أيضاً بصورة لأنواع السلوك التي سبقت أو تلت أنواع السلوك الأخرى.

لقد قسم فلاندرز حديث المعلم إلى (استجابة) أطلق عليها اسم "التأثير غير المباشر" ثم "المبادرة" التي أطلق عليها اسم "التأثير المباشر". وفي استقصاء باستخدام أصناف الرصد، درس فلاندرز أثر شكل التأثير غير المباشر لحديث المعلم مقارنة بحديث التأثير المباشر. وقد رصدت دروس في العلوم الاجتماعية والرياضيات كما قورنت الدروس ذات التعليم غير المباشر الأقصى مع دروس التعليم المباشر الأقصى حول التحصيل في وحدة تدريس جرى تعليمها خلال الرصد وكذلك حول الاتجاهات نحو المعلم. لقد وجد فلاندرز أن الفصول الدراسية ذات التعليم لأقصى المباشر كان أداؤها أفضل في قياسات الاتجاهات والتحصيل. وكان الاستنتاج هو إنه يمكن أن يتحسن عمل المعلمين في الصف، بتدريبهم على إجراء استخدام أكثر للأصناف غير المباشرة للسلوك الشفوي.

وفي دراسة صفية أخرى استخدم (Rollins, McCandless, & Thompson, 1974) رصداً منتظماً قبل وبعد فترة التدريب التي تعلم فيها المعلمون حيث تم تشجيعهم على استخدام إجراءات التعزيز الإيجابية. وأجريت حالات الرصد لتحديد المدى الذي يستخدم فيه المدرسون التعزيز الإيجابي وأثره على سلوك التلاميذ. وقد أنجز الرصد في فترات مدة كل منها خمس دقائق: خلال الفترة الأولى، جري عد أنواع سلوك التعزيز الإيجابية والسلبية للمعلم. وتم إجراء عد أنواع السلوك المشتت لدى التلاميذ خلال الفترة الثانية، أما خلال الفترة الثالثة. فقد تم رصد أنواع السلوك اليقظ للتلاميذ. لقد وجد الباحثون أن المعلمين الذين تدربوا قد أبدوا تعزيزات إيجابية بنسبة أربعة أضعاف لكل طالب كما استخدموا عقوبات أقل مما فعل المعلمون في صفوف المجموعة الضابطة. إن حالات رصد تلاميذ المعلمين المدربين قد أبدت نقصاً واضحاً في أنواع السلوك المشتت وزيادة في أنواع السلوك اليقظ. وعلاوة على ذلك، حصل التلاميذ في الصفوف التجريبية على معدل متوسط كسب أعظم في درجات اختبار القراءة مساو لأكثر من فترة ثمانية أشهر. وفي هذه الدراسة استخدم الرصد المنتظم بصورة فاعلة جداً لتوثيق التغيرات في أنواع السلوك الصفي التي ارتبطت بالتغيرات الواضحة في تحصيل التلاميذ.

تقييم الرصد المباشر Evaluation of Direct Observation

نستطيع في إطار الرصد المباشر الحصول على سجل بالسلوك الفعلي لدى الأشخاص في أوضاع طبيعية. وهذا أسلوب قيم بشكل خاص، إذ يمكن استخدامه مع الأطفال الصغار جداً ممن لا يستطيعون الاتصال عبر اللغة.

إن الرصد المباشر، على أية حال، إجراء باهظ التكاليف بسبب الزمن المطلوب للراصد. وعلاوة على ذلك، هناك دوماً احتمال أن يغير وجود الراصد سلوك الفرد. وقد تستخدم شاشات ذات رؤية باتجاه واحد في بعض الأوضاع للتغلب على هذه المشكلة. وقد وجد في العديد من الحالات أنه بعد رد الفعل الأولي، يعير الأشخاص المرصودون انتباهاً قليلاً للراصد، وخصوصاً من يعمل بصورة غير فضولية. ويتطلب الراصدون تدريباً واسعاً كي يصبحوا كفوءين في معرفة ما يرصدونه وكيفية التعامل مع حالات الرصد.

حالات الرصد المبتكرة / المخططة Contrived Observations

في حالات الرصد المبتكرة يرتب الباحث لرصد الأشخاص في أوضاع شبيهة بالأوضاع الحقيقية. وترتب الظروف كي يتم الحصول على أنواع السلوك المرغوب.

إن أحد أشكال الرصد المبتكر هو الاختبار الموقفي(*). والمثال التقليدي لهذا الاختبار – رغم أنه لم يسمى بهذا الاسم في ذلك الوقت – قد استخدم في سلسلة من الدراسات التي قام بها (Hartshorne & Nay, 1928) في استقصاء تربية الشخصية (CEI). لقد صممت هذه الاختبارات لاستخدامها في دراسة نمو السمات السلوكية كالنـــزاهة، والضبط الذاتي، والصدق، والتعاون. وقام هارتشورن وماي برصد الأطفال في نشاطاتهم المدرسية الروتينية إلا أنهما قاما أيضاً بترتيب بعض المواقف للتركيز على سلوك محدد. فمثلاً، عرضا على الأطفال مفردات واختبارات قراءة، وجمعا الاختبارات، ودون معرفة الأطفال عملا على نسخ أجوبتهم. وبعدئذ أعطي الأطفال مفاتيح الأجوبة وطلب منهم تسجيل الدرجات على أوراقهم الأصلية. والفرق بين الدرجات التي قدمها الأطفال والدرجات الحقيقية المستحصلة من الأوراق المستنسخة قد وفرت مقياساً للغش.

وفي اختبار آخر طلب إلى الأطفال وضع علامة في كل من عشر دوائر صفية وضعت بشكل غير منتظم مع إبقاء عيونهم مغمضة. وأشارت اختبارات سابقة تحت ظروف منعت اختلاس النظر، أن درجة، لأكثر من 13 علامة وضعت بالشكل الصحيح في 3 محاولات، كانت مستبعدة جداً. وهكذا فإن درجة أكثر من 13 تم تسجيلها كانت دليلاً على أن الطفل اختلس النظر.

لم يجد هارتشورن وماي عملياً أية علاقة بين الغش في مواقف مختلفة كما في الاختبار أو في ألعاب القوى. واستنتجا أن استجابات الأطفال كانت محددة بمواقفها – أي أن غش الطلبة اعتمد على نشاط محدد، وتدخل المعلم، ومواقف أخرى، وليس على سمة شخصية عامة.

(*) الاختبار الموقفي (Situational test): اختبار يوضع فيه الطالب في موقف عملي فعلي حيث يتم رصد تصرفاته وردود أفعاله– (المراجع)

الخلاصة SUMMARY

إن إحدى الواجبات المهمة لدى الباحثين في العلوم السلوكية، تكون في اختيار و/أو تطوير أدوات قياس موثوقة بهدف التحديد الكمي للمعلومات البحثية. والاختبارات في البحوث التربوية، هي أدوات قياس واسعة الاستخدام. ويعرف الاختبار بكونه مجموعة مثيرات تعرض على شخص بهدف الحصول على استجابات يمكن على أساسها تخصيص درجة عددية له. وتعد اختبارات التحصيل أمثلة رئيسية لمثل هذا النوع من الأدوات. وهناك عدة أنواع من اختبارات التحصيل المتيسرة لتأمين متوسطات يمكن استخدامها كأساس للمقارنة. أما اختبار الأداء فيقيس ما يمكن أن يفعله المرء وليس ما يعرفه. واختبارات الاستعداد أدوات لتقييم الطاقات الفعلية وغير الفعلية للشخص. وتصمم استبيانات الشخصية لقياس السمات الشخصية والأداء النموذجي للفرد.

إن مقاييس الاتجاهات، أدوات لقياس معتقدات ومشاعر وردود أفعال الأشخاص إزاء أشياء معينة. والأنواع الرئيسية لمقاييس الاتجاهات هي مقاييس ليكرت ومقاييس ثيرستون وتمايز المعاني.

وتحدد مقاييس التقدير درجات للتقييمات التي يجريها الأشخاص أو الراصدون فيما يخص مفهوماً معيناً أو شيئاً ذا اهتمام وتستخدم بعدئذ هذه الدرجات لقياس المفاهيم البنائية الضمنية. ومن بين أنواع مقاييس التقدير هناك المقياس البياني والمقياس التصنيفي ومقاييس التقدير المقارن.

أما أساليب القياس الاجتماعي فهي أدوات لتقييم مواقف الأشخاص بين أقرانهم. ويمكن من خلال هذه الأساليب تحديد الأفراد المشهورين في المجموعات (النجوم) والمنعزلين والزمر.

لقد تم تطوير عدد من الإجراءات الخاصة للرصد المنظم لسلوك الأفراد كطريقتي رصد مباشر وأخرى لرصد مبتكر.

مفاهيم أساسية Key Concepts

achievement test	اختبار إنجاز / تحصيل
aptitude test	اختبار استعداد
attitude scale	مقياس اتجاهات
contrived observation	رصد مبتكر / مخطط
direct observation	رصد مباشر
error of central tendency	خطأ النزعة المركزية
error of severity	خطأ الصرامة / القسوة
generosity error	خطأ الكرم

halo effect	تأثير / ظاهرة الهالة
inventorv	استبيان
Likert scale	مقياس ليكرت
performance assessment	تقييم الأداء
projective technique	الطريقة الاسقاطية
scale	مقياس / سلّم قياس
scholastic aptitude test	اختبار الاستعداد الدراسي
semantic differential scale	سلّم / مقياس تمايز المعاني
sociogram	مبيان اجتماعي
standardized test	اختبار مقنن
summated rating scale	مقياس التقدير الجمعي
test	اختبار
Thurstone scale	مقياس ثيرستون

تمارين EXERCISES

1. ما معنى مصطلح (مقنّن) عندما يطبق على أدوات القياس؟
2. ما الفرق بين مقاييس التقدير المقارن والمقاييس البيانية والتصنيفية.
3. أدرج بعض مصادر التحيز المألوفة في مقاييس التقدير.
4. ما نوع الأدوات التي قد يختارها الباحث بغية الحصول على بيانات حول كل ما يأتي؟
   أ. كيف يشعر طلبة الكلية بشأن تشريع إباحة المارجوانا.
   ب. إمكانية طلبة السنة النهائية في كلية صغيرة للنجاح في كلية للدراسات العليا.
   جـ. معرفة ما إذا كان بوسع طلبة الثانوية في درس الكيمياء تحليل مركب كيمياوي مجهول.
   د. مدى مقارنة الطلبة في مدرسة براون الابتدائية حسب المتوسط القومي في مهارات القراءة.
   هـ. أنماط الإرشاد المفضلة لمجموعة من طلبة السنة الأولى في الكلية.
   و. مدى أداء طلبة في مسابقة الخطابة.
   ز. الولد الأكثر شعبية في صف السنة الأولى للمعلمة سميث.
   حـ. القدرات اللفظية وغير اللفظية لطالب يعاني من عجز في الانتباه.

ط. المدى الذي يستخدم فيه مدرسو الابتدائية التعزيزات السلبية في الصف وأثر ذلك التعزيز على سلوك الطلبة.
ي. المشكلات التي تواجه طلبة الأقليات خلال السنة الأولى في جامعة بحوث كبيرة.
ك. كيف يشعر الآباء في نظام مدرسي إزاء نقل الصف السادس من الابتدائية إلى المتوسطة.

5. ما الأداة الرئيسة التي تستخدم في مقياس تمايز المعاني؟

6. ما هي بعض الإجراءات لزيادة دقة أساليب الرصد المباشر؟

7. أنشئ مقياس ليكرت ذي الخمسة بنود لقياس اتجاهات المعلمين إزاء تدريس لغة أجنبية في المدرسة الابتدائية.

8. ما هي الميزات الرئيسة لمقياس ليكرت بالنسبة لمقياس ثيرستون؟

9. أنشئ مقياس تقدير بياني من خمسة بنود يكون ذا فائدة لتقييم برنامج بحثي.

10. تمثل البيانات التالية تقديرات خصصت لبند واحد من 200 حكم/خبير في إجراء ثيرستون. ما هي القيمة التي ستخصص لهذا البند؟

الصنف	عدد الحكام
11	صفر
10	صفر
9	صفر
8	30
7	50
6	60
5	34
4	16
3	10
2	صفر
1	صفر
	200

11. فيما يلي قيم مقياس لخمسة بنود وضع إشارات عليها طالبان في مقياس ثيرستون احسب درجتي اتجاهاتهما، وفسر الدرجات في إطار مدى التفضيل للاتجاهات والتحديد الواضح لها.

أ	ب
5.5	10.5
4.6	8.2
4.1	7.1
3.8	3.9
3.1	2.8

الأجوبة ANSWERS

1. يشير (مقنن) إلى أدوات اشتقت من أجلها متوسطات المقارنة وتحقق ثباتها وصدقها كما وصفت توجيهات لتطبيقها ووضع الدرجات.

2. في الحكم على شخص في مقاييس بيانية وتصنيفية، لا يقوم المقدرون بإجراء مقارنة مباشرة للشخص مع الآخرين. وفي الحكم على شخص في مقياس تقدير مقارن يجب أن يكون لدى المقدر المصنف معرفة بالجماعة التي يقارن بها الشخص.

3. قد يكون المقدرون أقل موضوعية في الحكم على الأشخاص عندما يتأثرون بنـــزعات مثل تأثير الهالة، أو خطأ الكرم، أو خطأ الصرامة / القسوة، أو خطأ النـــزعة المركزية.

4. أ. مقياس اتجاهات.
   ب. اختبار الاستعداد (مجموعة).
   جـ. اختبار الأداء.
   د. اختبار الإنجاز / التحصيل.
   هـ. استبيان.
   و. مقياس تقدير (اختبار أداء).
   ز. أسلوب القياس الاجتماعي.
   ح. اختبار استعداد أو ذكاء (فردي).
   ط. رصد مباشر.
   ي. استبيان.
   ك. مقياس الاتجاهات.

5. يستخدم مقياس تمايز المعاني لقياس المعنى الدلالي التضميني الذي يلحقه الشخص بموضوع معين.

6. يجب أن تكون أنواع السلوك المطلوب ملاحظتها محددة. ويجب تحديد أنواع السلوك التي تقع ضمن صنف معين، كما يجب تطوير نظام للتكميم وتدريب الراصدين لتنفيذ الرصد وفقاً لهذا الإجراء القائم.

7. ستختلف الأجوبة.

8. الفائدة الرئيسية هي أن مقياس ليكرت أسهل بناءاً ومن المحتمل أيضا أن يكون أكثر ثقة من مقياس ثيرستون الذي يحتوي على العدد ذاته من البنود.

9. ستختلف الأجوبة.

10. 6.17

11. مع وسط يبلغ 4.2 ووسيط 4.1 يبدي أ اتجاهات مفضلة معتدلة. كما أن خياراته قريبة من بعضها في قيم المقياس مما يوضح الثبات. أما وسط ب البالغ 6.5 ووسيطه البالغ 7.1 فيوضح موقفاً غير مفضل قليلاً. أما الانتشار الواسع لقيم المقياس فإنها توضح اتجاها غير محدد بشكل واضح.

REFERENCES المصادر

Flanders, N.A. (1965). Teacher influence, pupil attitudes, and achievement. (U.S. Office of Education Cooperative Research Monograph). Washington, D.C.: U.S. Government Printing Office.

Flanders, N.A. (1970). Analyzing teaching behavior. Boston: Addison-Wesley.

Gronlund, N.E., and Linn, R.L. (1990). Measurement and evaluation in teaching (6th ed.). New York: Macmillan.

Hart, D. (1994). Authentic assessment: A handbook for educators. Menlo Park, CA: Addison-Wesley.

Hartshorne, H., and May, M.A. (1928). Studies in the nature of character: Studies in deceit (Vol. 1). New York: Macmillan.

Likert, R. (1932). A technique for the measurement of attitudes. Archives of Psychology, No. 140.

Mehrens, W.A., and Lehmann, I.J. (1991). Measurement and evaluation in education and psychology (4th ed.). Fort Worth, TX: Holt, Rinehart and

Winston.

Mueller, D.J. (1986). Measuring social attitudes: A handbook for researchers and practitioners. New York: Teachers College Press.

Osgood, C.E., Suci, G.J., and Tannenbaum, P.H. (1967). The measurement of meaning. Urbana: University of Illinois Press.

Rollins, H.A., McCandless, B.R., and Thompson, M. (1974). Project success environment: An extended application of contingency management in inner-city schools. Journal of Educational Psychology, 66, 167-178.

Sax, G. (1989). Principles of educational and psychological measurement and evaluation (3d ed.). Belmont, CA: Wadsworth.

Stiggins, R.J. (1992). High quality classroom assessment: What does it really mean? Educational Measurement, 11, 35-39.

Thurstone, L., and Chave, E. (1929). The measurement of attitude. Chicago: University of Chicago Press.

Urban, J. (1943). Behavior changes resulting from a study of communicable diseases. In Teachers College contributions to education (No. 896). New York: Teachers College Press, Columbia University.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. Phi Delta Kappan, 71, 703-704.

Wiggins, G. (1993). Assessment: Authenticity, context, and validity. Phi Delta Kappan, 7,5, 200-214.



الصدق / الصحة والثبات / الاستقرار
VALIDITY AND RELIABILITY

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون الطالب قادراً على أن:

1. يميز بين الصدق والثبات / الاستقرار.
2. يصف ثلاثة أنواع رئيسية للصدق والطرق المستخدمة في تقييمها.
3. يختار نوع الصدق المناسب لأهداف القياس المختلفة.
4. يميز بين الصدق التقاربي وصدق التمييز.
5. يشرح العلاقة بين الثبات ومفهوم أخطاء القياس العشوائية.
6. يذكر مصادر الخطأ المختلفة في القياسات التربوية والنفسية.
7. يصف الإجراءات المختلفة (إعادة الاختبار، الصور المتكافئة، التجزئة النصفية، كودر – ريتشاردسون وأمور أخرى) لتقييم ثبات القياس.
8. يحسب معاملات الثبات من بيانات معينة.
9. يعرف الثبات الداخلي للمراقبين ويوضح كيفية حسابه.
10. يطبق معادلة سبيرمان – براون لتحديد أثر إطالة الاختبار على ثبات الاختبار.
11. يشرح العوامل المؤثرة على حجم معامل الثبات.
12. يفسر الخطأ المعياري للقياس ويشرح علاقته بثبات / استقرار الاختبار.

13. يحسب الخطأ المعياري للقياس ويفسر مدى الدرجات.
14. يحسب المؤشرات لإيضاح ثبات الاختبار مرجعي الإتقان.

يعتمد البحث دائماً على القياس. وهناك خاصتان مهمتان ينبغي أن تحوز عليها كل أداة قياس وهما: الصدق والثبات. ويشير الصدق إلى المدى الذي تقيس فيه الأداة ما يراد قياسه. أما الثبات، من ناحية أخرى، فهو المدى الذي تظل فيه أداة القياس ثابتة في قياس ما تقيس. وبصورة محددة، فإن الصدق والثبات يشيران إلى المعلومات التي تنتجها أداة القياس وليس إلى الأداة ذاتها.

ولغرض السهولة، نتحدث عن صدق وثبات اختبار معين، إلا إن الأكثر دقة هو الحديث عن صدق وثبات درجات الاختبار. ولما كان الاختبار يتصف بالسكون فإن النتائج (الدرجات) تتغير حسب الوضع. فيجب على الباحث أن يستقصي صدق وثبات الأدوات المستخدمة في أية دراسة، ويجب عليه تضمين هذه المعلومات في تقرير البحث. وإن لم يتم الحصول على بيانات الباحث بأدوات صادقة وموثوقة، فستكون لدى المرء ثقة ضئيلة بالنتائج أو بالاستنتاجات القائمة على تلك النتائج.

ويعدّ دليل الصدق والثبات مهماً بشكل خاص في البحوث التربوية لأن أغلب القياسات التي نسعى إليها في هذا المجال يتم الحصول عليها بشكل غير مباشر. فالمرء بحاجة إلى تقييم المدى الذي تقيس فيه أداة قياس تربوية أو نفسية ما يراد قياسه بشكل موثوق.

الصدق VALIDITY

الصدق، كما أوضحنا أعلاه، يتعلق عموماً بالمدى الذي تقيس فيه أداة معينة ما يفترض أنها تقيسه. ويعرف الصدق، بشكل أكثر تحديدا، بكونه الصحة والدلالة الهادفة والفائدة للاستدلالات المحددة الناجمة عن درجات الاختبار. ومن الضروري، بصورة مطلقة، أن يأخذ الباحث مسألة الصدق بنظر الاعتبار. فأدوات الاختبار النفسي والتربوي مصممة بهدف تقييم المفاهيم البنائية مثل التحصيل، والذكاء، والإبداع، والاستعداد، والاتجاهات، والدافعية وما يماثلها. وعلى أية حال، ليس هناك أدوات مباشرة لقياس هذه المفاهيم البنائية كالتي توجد في العلوم المادية لقياس خواص مثل الطول، والحجم، والوزن. فلابد للباحثين من تطوير أدوات غير مباشرة لقياس السمات المعقدة. وتنطوي هذه الأدوات غير المباشرة على اختبارات ومقاييس مكونة من عدد من المهام المنتخبة كي تعمل كمؤشرات للمفاهيم البنائية المعقدة. وقد يسأل المرء عن كفاءة مثل هذه الإجراءات غير المباشرة في قياس ما يفترض قياسه. وعلى الباحثين أن يطرحوا أسئلة مثل: هل يسمح هذا الاختبار

الصفي للمعلم أن يستدل على المدى الذي حقق فيه الطلبة أهداف المقرر؟ هل يقيس هذا الاختبار حقاً دافعية التحصيل؟ هل يقيس هذا الاختبار صفات أخرى كذلك؟ هل بوسع اختبار الإبداع هذا أن يفصل الأشخاص المبدعين جداً عن الأشخاص الأقل إبداعاً؟ هل بوسع المرء إقامة تنبؤات مفيدة استناداً إلى الدرجات في اختبار الاستعداد؟ هل هو أداة مناسبة يمكن استخدامها مع جميع الطلبة أم هل ينبغي استخدامها فقط مع مجموعات معينة؟ فمثلاً، ما مدى ملاءمة اختبارات الاستعداد للتنبؤ بالتحصيل الأكاديمي لدى طلبة الأقليات؟ إن جميع هذه الأسئلة تتعلق بدلالة وفائدة الاستنتاجات التي تستمد من الدرجات – أي صدق الاختبار. أما عملية جمع الأدلة للإجابة على مثل هذه الأسئلة فتدعى بالصدق.

إن الصدق يكون دوماً خاصاً بهدف معين تستخدم الأداة من أجله. فالاختبار الذي يتمتع بالصدق في وضع ما ومن أجل هدف ما، قد لا يتمتع بالصدق في وضع مختلف أو من أجل هدف آخر. فمثلاً قد يكون لاختبار مقنن في التاريخ يؤكد على الفهم والتفسير، صدق بالنسبة لمدرس التاريخ الذي يؤكد على هذه الأهداف الإدراكية المعرفية في الصف لكنه لن يكون صادقاً بالنسبة لمدرس أخر يؤكد على تعلم التواريخ ومعرفة الوقائع. فالهدف الذي يستخدم لأجله الاختبار هو كذلك عامل كبير في الصدق. فالاختبار المقنن لمادة الكيمياء قد يستخدم لقياس تحصيل نهاية السنة في مادة الكيمياء. وفي هذه الحالة يكون لدى المرء أسئلة حول ما يقيسه الاختبار ومدى جودته في القيام بذلك. وقد يستخدم نفس الاختبار للتنبؤ بالتحصيل في مادة الكيمياء في الكلية. ففي هذه الحالة قد يهتم المرء بالمدى الذي يمكن للاختبار أن يتنبأ فيه، بالتحصيل المستقبلي. فالأهداف المختلفة للاختبارات تتطلب أنواعاً مختلفة من الشواهد لدعم صدق ذلك الاستخدام المعين.

ورغم إن الصدق مفهوم موحد إلا إن هناك أنواعاً من الأدلة يمكن جمعها لدعم الاستدلالات المستمدة من درجات أداة القياس. فالمعايير المعدة من لجنة مشتركة من الجمعية النفسية الأمريكية (APA) وجمعية البحوث التربوية الأمريكية (AERA) والمجلس القومي للقياس في التربية (NCME) تصنف أدوات جمع أدلة الصدق في الأصناف التالية: دليل صدق المحتوى، ودليل صدق المعيار، ودليل صدق المفهوم البنائي (الجمعية النفسية الأمريكية، 1985).

الدليل المتعلق بالمحتوى Content - Related Evidence

تستخدم الاختبارات الصفية عموماً لهدف تقييم معرفة ومهارات الطلبة في محتوى محدد. والطريقة المثالية لتحقيق ذلك تكون باستخدام اختبار يتضمن جميع الأسئلة التي يمكن أن تسأل حول المحتوى. ومن الواضح أن مثل هذا الإجراء غير عملي. أما البديل المفيد فهو

اختيار عينة من مجال المحتوى الكلي واستخدام هذه العينة كأساس لاستدلالات حول معرفة الطلبة بشمولية المحتوى. ونظراً لأن الاستدلالات تستمد أساساً من عينة من الأسئلة فقط، فإن من المهم أن تكون العينة ممثلة للمحتوى برمته - أي أن تكون عينة صادقة.

وهذا يقودنا إلى مسألة الدليل المرتبط بالمحتوى. ويبين هذا النوع من الدليل المدى الذي تكون فيه عينة بنود الاختبار ممثلة لمجتمع إحصائي محدد، أو مجال المحتوى. وتجمع الأدلة عن طريق اختبار دقيق وناقد من قبل حكام خبراء بمحتوى الاختبار لتحديد العلاقة بين الاختبار والمجتمع المحدد. وينفذ هذا الإجراء من وجهة نظر الاستخدام الخاص لنتائج الاختبار.

وبالطبع، فإن المجتمع الإحصائي لمثل هذا المحتوى نظري. ففي الإعداد الفعلي للاختبار ينبغي على المرء إعداد خطة للموضوعات والمهارات والقدرات التي تكون مجال المحتوى المراد قياسه جنباً إلى جنب مع إيضاح أهمية كل منها. ويكتب عدد كبير من بنود الاختبار باستخدام هذه الخطة كمرشد. ومن كل صنف في الخطة يمكن أن تستمد بنود الاختبار بشكل عشوائي على أن يعكس عدد الوحدات الوزن النسبي لذلك الصنف في الكل. فينبغي أن تكون عينة البنود الناتجة ممثلة لمجتمع المحتوى. فمثلاً، قد يتحدد مجتمع المحتوى لاختبار حول الحرب الأهلية بكونه معرفة وفهم موضوعات مثل الأسباب والاستراتيجيات العسكرية والحملات، والشخصيات المهمة والأثر الاقتصادي، والآثار على التاريخ اللاحق، وما إلى ذلك. ولا ينبغي لبنود الاختبار المكتوبة على كل من هذه الموضوعات أن تقيس المعرفة بالموضوع فحسب، بل كذلك الفهم والتفسير والتحليل وأية أهداف إدراكية / معرفية أخرى مؤكدة في المقرر. وينبغي لعدد البنود التي تغطي كل موضوع وكل نوع من الأهداف، أن يعكس التأكيد على ذلك الموضوع وذلك الهدف في المقرر كله. أفترض إن مدرس الأدب الإنكليزي قد أكد على فهم أفكار مؤلفين منتقين وعلى صلة بتلك الأفكار بالقرن العشرين. فإذا احتوى اختبار هذا المدرس على مزاوجة أسماء المألوفين مع أعمالهم واستذكار تواريخ ميلادهم فإن المدرس يخفق في الحصول على عينة ممثلة لنطاق المحتوى، ويكون للاختبار صدق ضئيل في هدفه المقصود وهو قياس فهم أفكار المؤلفين.

ولا يعبر عادة عن الدليل المرتبط بالمحتوى بصورة عددية. وجمع مثل هذا الدليل يستند أساساً وبالضرورة إلى الحكم، وإن مثل هذا الحكم يجب أن يتم بشكل منفصل لكل هدف. وهذا ينطوي على اختبار دقيق وناقد لتحديد ما إذا كان المحتوى والأهداف المقاسة بالاختبار ممثلة لتلك التي يتكون منها مجال المحتوى. وينبغي أن يحدد المرء ما إذا كانت بنود الاختبار تمثل المقرر والأهداف كما هي موضحة في أدلة المناهج والمفردات والنصوص. ولغرض الحصول على تقييم خارجي لصدق المحتوى، ينبغي على واضع

الاختبار أن يطلب من عدد من المعلمين، أو خبراء آخرين، فحص محتوى الاختبار بشكل منتظم وتقييم مدى صلته بمجتمع محدد. فإن أتفق الجميع أن بنود الاختبار تمثل مجال المحتوى بشكل مناسب، فإنه يمكن القول أنه يتمتع بصدق المحتوى. ويجب التثبت أيضاً من أن الاختبار خال من تأثير العوامل التي لا علاقة لها بهدف القياس. فمثلاً، قد يود المرء أن يكون لسرعة القراءة والمفردات أقل تأثير ممكن على الأداء في اختبار الرياضيات. فوجود مثل هذه العوامل في اختبار الرياضيات، سوف يقلل من صدقه لأن الاختبار سوف يقيس شيئاً آخر غير ما أريد قياسه. ويعد الدليل الخاص بالمحتوى ذا أهمية خاصة لتقييم صدق اختبارات التحصيل.

ويجب على الباحث دوماً أن يستخلص الدليل المتعلق بالمحتوى لصدق أي اختبار يُبنى ذاتياً أو اختبار تحصيل مقنن سيستخدم في الدراسة. فناشرو الاختبارات يقدمون عموماً دليلاً واسعاً للصدق من هذا النوع. وعلى أية حال، ينبغي التأكيد مرة أخرى أنه قد يكون لاختبار التحصيل صدق عال بالنسبة لمن وضعه وقد لا يكون له صدق لدى مستخدم آخر، قد يحدد مجتمع المحتوى بطريقة مختلفة. وفقط، مستخدم اختبار التحصيل يمكن له أن يحكم بشكل نهائي على صدقه حسب هدفه.

الدليل المتعلق بالمعيار Criterion - Related Evidence

يبين الدليل المتعلق بالمعيار، المدى الذي ترتبط به درجات أداة القياس مع متغير خارجي مستقل (معيار) يعتقد أنه يقيس بشكل مباشر السلوك أو الخاصة المعنية. فحين يستقصي المرء العلاقة بين درجات اختبار الاستعداد / التحصيل الدراسي (SAT) ومعدل نقاط الكلية (GPA) فإنه يستخلص دليلاً متعلقاً بالمعيار لأجل صدق الاختبار. إن المدى التي ترتبط به درجات اختبار الاستعداد بالنجاح في الكلية كما هو مقاس حسب (GPA) هو المدى الذي يكون فيه لاختبار الاستعداد صدق لأجل تنبؤ (GPA).

وكما يوضح الاسم، فإن التأكيد في هذا النوع من الدليل يكون على المعيار وإجراءات القياس المستخدمة للحصول على درجات المعيار. إن اختيار المعيار مهم لنجاح هذا النوع من الاستقصاء. وهناك عدة خواص ينبغي أن يحوزها مقياس المعيار. ولعل أهمها هو "وثاقة الصلة بالموضوع". فالمرء يجب أن يحكم ما إذا كان المعيار المختار يمثل فعلاً الأداء الناجح للسلوك المعني. فإذا لم يعكس المعيار الخاصة قيد الدراسة فإنه لن تكون ثمة جدوى من استخدامه كأساس لتحقيق صدق على أداة أخرى. ويعتبر GPA مقياساً ذا صلة بالنجاح في الكلية، ويتم اختياره، عموماً، كمعيار لتحقيق صدق اختبارات الاستعداد التي تجري لاختيار المتقدمين إلى الكلية. ولتحقيق صدق اختبار مصمم لاختيار بائعين فقد يكون المعيار المناسب هو قيمة المبيعات بالدولار التي جرت في زمن محدد. وقد

يصعب في بعض الحالات إيجاد معايير مناسبة. فمثلاً، كان من الصعب تحديد معايير تستخدم لتحقيق صدق مقاييس تستخدم للتنبؤ بفاعلية المعلم. فمع عدم وجود، وصف متفق عليه لفاعلية المعلم أو طريقة قياس ناجحة لذلك المتغير، فإنه يستحيل عملياً تحقيق صدق أية أداة مصممة لتحديد معلمين مرجوين مرشحين.

والخاصة الثانية لمقياس المعيار هي وجوب الوثوق به. وهذا يعني أنه ينبغي للمعيار أن يكون مقياساً ثابتاً للخاصة على مدار الزمن أو من موقف لآخر. فإن لم يكن المعيار متساوقاً / مستقراً فلا يمكن أن نتوقع ارتباطه بشكل ثابت بأية أداة تنبؤ.

وينبغي أن يكون المعيار "متحرراً من التحيز"، أي لا ينبغي أن يتأثر وضع درجات مقياس المعيار بأية عوامل غير الأداء الفعلي الخاص بالمعيار. افترض أن تقدير المراقب هو المعيار المستخدم لتحقيق صدق اختبار لأجل اختيار مرشحين لعمل معين. فلو سمح المراقب للرأي العام حول الشخص أو أي عامل آخر، غير الأداء الفعلي بالتأثير على التقدير. فستكون درجة المعيار متحيزة. ولغرض تجنب التحيز عندما يكون المعيار هو التقدير، ينبغي على المرء إعطاء تعليمات واضحة عن الخواص المراد تقديرها وكيفية إجراء التقدير. وكلما كان إجراء التقدير موضوعياً كان التحيز قليلاً في المعيار. وثمة مصدر محتمل آخر للتحيز في المعيار هو "التلوث / الإفساد". فالتلوث يحدث عندما تكون درجة الفرد حسب المعيار متأثرة بمعرفة واضع التقدير بأداة التنبؤ لتقدير الفرد. فمثلاً، افترض أن أحداً لديه اختبار الاستعداد الفني يراد تحقيق صدقها باستخدام الدرجات في دروس الفن كمعيار. فإذا كان المدرسون الذين يقدرون درجات أعمال الطلبة مدركين لدرجات الطلبة في اختبار الاستعداد فإن مثل هذا الإدراك قد يؤثر على تقييم المدرسين للطلبة. ويمكن حجب تلوث / إفساد المعيار بعدم السماح للشخص الذي يعطي الدرجات أو التقديرات أن يرى درجات أداة التنبؤ.

وحالما يتم تحديد المعيار الخارجي، تجمع البيانات التجريبية بغية تقييم العلاقة بين الدرجات على أداة قياس (X) وعلى معيار (Y). وتعرض الأداة المراد تحقيق صدقها على مجموعة من الأفراد يمثلون أولئك الذين سيستخدم القياس عليهم. وتوضع درجات الأفراد حسب أداة التنبؤ (X) جانباً، ولا تستخدم لاتخاذ أية أحكام قد تؤثر على الأحداث التالية لهذه المجموعة المحددة بغية تجنب إفساد درجات المعيار. وعندما تتوفر بيانات المعيار (Y) في وقت تال نعود للاختبارات الأصلية ويتم دراسة الترابط بين درجات الاختبارات مع درجات المعيار. إن معامل الارتباط الحاصل بين هاتين المجموعتين من القياسات يدعى "معامل ارتباط الصدق" (r_{xy}) وهو يوضح مدى دقة تنبؤ درجات الاختبار (X) بالمعيار (Y)، وكلما كان (r_{xy}) كبيراً كان تنبؤ الاختبار أدق.

إن مثالاً عن هذا الإجراء حدث في تحقيق صدق اختبار الاستعداد الدراسي (SAT).

ففي دراسات متعددة جري اختبار (SAT) على عدد كبير من طلبة الصفوف المنتهية في المدارس، وتم تنحية الاختبارات جانباً حتى أكمل الطلبة السنة الأولى من الكلية. وحينذاك تم دراسة الترابط بين درجات SAT ومعدل نقط GPA السنة الأولى (المعيار) بغية الحصول على معامل صدق الاختبار. ونظراً لأنه تبين في دراسات متكررة أن SAT يمتلك صدقاً مفيداً متعلقا بالمعيار، فإنه يستخدم الآن بشكل روتيني للتنبؤ بالأداء في الكلية. فطلبة الثانوية عادة ما يأخذون اختبار SAT خلال سنتهم المنتهية ويقدمون الدرجات إلى الكليات. ويقوم مسؤولو القبول في الكلية المدركون لصدق SAT التنبؤي بفحص الدرجات واتخاذ قرارات القبول استناداً إلى درجات SAT، ولو جزئياً على الأقل. فكلما كانت درجات SAT عالية كان احتمال النجاح في الكلية أعلى. وفي أغلب الحالات يستند التنبؤ على ارتباط درجات SAT مع بعض مقاييس التحصيل في المدرسة الثانوية مثل المرتبة في السنة الثانوية الأخيرة. إن ارتباط درجة SAT ومرتبة السنة الثانوية الأخيرة يتنبأ بمعدل النقط GPA في الكلية بشكل أكثر دقة من أي من المقياسيين فحسب.

إن إحدى الطرق الجيدة لعرض الدليل المتعلق بالمعيار لصدق الاختبار تكون استخدام "جدول التوقع". فجدول التوقع هو شبكة ذات طريقين تدرج فيها درجات الاختبار (أداة التنبؤ) على محور عمودي وتدرج درجات المعيار على محور أفقي. فدرجات أداة التنبؤ والمعيار تجمّع في أصناف، ويبين الجدول لكل من أصناف أداة التنبؤ نسبة الأشخاص الذي يقعون في كل صنف للمعيار. ويعرض جدول التوقع احتمال نتائج المعيار المختلفة للأشخاص الذين تعطى درجات أداة التنبؤ لهم أو لمجال من هذه الدرجات. إن مثال التوقع الافتراضي على البيانات المتراكمة مبين في الجدول 8.1.

الجدول 8.1 مثال عن جدول التوقع

الدرجة الكلية في SAT	معدل نقاط السنة الأولى (GPA)		
	0.0-1.9	2.0-2.9	3.0-4.0
1350 أو أعلى	0	6	94
1230 - 1349	2	40	58
1140 - 1229	9	56	35
1050 - 1139	20	60	20
970 - 1049	35	50	15
900 - 969	50	45	5
840 - 899	52	44	4
780 - 839	56	42	2
779 أو أقل	65	33	2

يمكن للمرء باستخدام جدول التوقع تقييم كل GPA محتمل للطالب بالاستناد إلى الدرجة الكلية في SAT. مثلاً، يبين الجدول أن 35% من الطلبة الذين حصلوا على درجة كلية (1200) في SAT قد حصلوا على معدل GPA في السنة الأولى مقداره (0.3) أو أعلى، وأن (56%) كان لهم GPA بين (2.0) و (2.9)، وأن (9%) كان لهم GPA (1.9) أو أقل. وهكذا بالنسبة للطالب الذي حصل على (1200) في اختبار SAT تكون الفرص هي (35) من (100) بأن GPA لديه سيكون (3.0) أو أعلى، وفي (56) من (100) سيكون بين (2.0) و (3.0)، وفي (9) من (100) سيكون (1.9) أو أقل. وبدمج الأصناف يمكن القول إن الفرص هي 91 من 100 (35+56) في أن يحصل الطالب على GPA مقداره (2.0) أو أعلى. وفي الطرف الآخر بالنسبة للطالب الذي حصل على درجة كلية في SAT مقدارها (800) يمكن القول إن احتمال حصوله على تقدير (3.0) أو أعلى هو فقط (2) من (100)، بينما فرص الحصول على GPA من (1.9) أو أقل هي (56) من (100).

ويعد الدليل المتعلق بالمعيار جوهرياً للاختبارات المستخدمة لأهداف الاختيار والتصنيف. وقبل استخدام اختبار معين للاختيار يجب أن يكون لدينا دليل على أن بوسع الاختبار أن يتنبأ فعلاً بالأداء حسب معيار محدد. والسؤال الأساسي هو: ما مدى الدقة التي يمكن التنبؤ فيها بالأداء المعياري (الأداء وفقاً للمعيار) من درجات الاختبار؟ وسواء كانت هذه عالية أو متدنية، مفيدة أم غير مفيدة، فإن الأمر يعتمد على السياق الذي يستخدم فيه الاختبار. وقد يكون معامل الارتباط البالغ (0.40) مفيداً جداً في الحالات التي لا تتوفر فيها أداة سابقة للتنبؤ. وفي حالات أخرى قد يعتبر معامل الارتباط البالغ (0.65) متدنياً وغير مقنع إذا توفرت أدوات تنبؤ أخرى وكانت لها علاقة أعلى بالمعيار. وبشكل عام فإن للأداة صدقاً "مفيداً" كوسيلة للاختيار إذا بين الدليل أن معامل ارتباط الأداة بالمعيار أعلى من الأدوات المنافسة. ويتطلب تجميع الدليل المتعلق بالمعيار من أجل صدق الاختبار، الوقت والصبر. وفي بعض الحالات يكون من الضروري الانتظار لعدة سنوات لتحديد ما إذا كان الأداء في مقياس معين مفيداً للتنبؤ بالنجاح في المعيار.

ويمكن التمييز بين تصميمين للحصول على الدليل المتعلق بالمعيار: الدراسات التنبؤية والدراسات المتزامنة. ويهتم كلاهما بالعلاقة التجريبية بين درجات الاختبار والمعيار، غير أن فارقاً يقوم بينهما على أساس الزمن عندما تجمع بيانات المعيار. فالدراسة التنبؤية تجمع المعلومات حول معامل الارتباط بين درجات الاختبار والمعيار الذي يحدث في وقت مستقبلي. وتجمع الدراسة المتزامنة المعلومات حول معامل الارتباط بين درجات الاختبار ومقياس المعيار المتيسر في ذات الوقت. فمثلاً، قد تنظر الدراسة المتزامنة إلى العلاقة بين درجات اختبار القراءة الذي يجري في نهاية الصف الرابع، ودرجات المعلمين للأفراد ذاتهم في مادة القراءة للصف الرابع، بيد أن الدراسة التنبؤية تقوم بدارسة معامل الارتباط بين

درجات الاختبار ذاتها ودرجات الطلبة التي يتم الحصول عليها في مادة القراءة للصف الخامس. ويفضل الدليل التنبؤي عموماً في تحقيق صدق اختيار الاختبارات في التربية أو الصناعة، بينما تفضل الدراسات المتزامنة عموماً في اختبارات التحصيل والاختبارات المصممة لقياس المفاهيم البنائية أو الاختبارات المستخدمة في المصادقة / الإشهاد أو التشخيص.

ونظراً لأن معامل الصدق هو معامل ارتباط، فإن حجمه سوف يتأثر بالعوامل نفسها التي تؤثر على أي معامل ارتباط - أي خطية العلاقة بين الاختبار والمعيار ومدى الفروقات الشخصية في المجموعة.

الدليل المتعلق بالمفهوم البنائي

Construct - Related Evidence

يركز الدليل المتعلق بالمفهوم البنائي على درجات الاختبار كمقياس لسمة نفسية أو مفهوم بنائي. لنتذكر من الفصل الأول أن المصطلح (Construct) يشير إلى شيء ما لا يقاس بشكل مباشر بل بما يفسر أثاراً يمكن ملاحظتها. فالمفهوم البنائي "النضج الاجتماعي" قد (بُني) لتفسير أنماط سلوك مرصود. فالنضج الاجتماعي، لا يمكن قياسه بشكل مباشر، إلا إن العديد من أنواع السلوك التي نعتقد أنها مظاهر هذا المفهوم البنائي يمكن وصفها وقياسها. وأن مجموع هذه القياسات يمكن أن يمدنا بقياس غير مباشر للمفهوم البنائي المجرد وهو "النضج الاجتماعي". وبعض الأمثلة المألوفة الأخرى للمفاهيم البنائية هي القلق، والذكاء، والدافعية، والقدرة الاستدلالية، والتفكير الناقد، والاستعداد في عدة مجالات، والاستيعاب القرائي، ومفهوم الذات.

لقد دمج الناس أو أنشاؤا عبر التاريخ تجريدات أكثر تعقيداً من مفاهيمهم. فتماماً كما يجمع طفل قطعاً في لعبة غير متقنة يسميها "حصاناً" أو "رجلاً" كذلك يبتدع الناس مفاهيم بنائية بدمج مفاهيم ومفاهيم بنائية أقل تعقيداً في أنماط هادفة. وجاء حافز تحقيق صدق المفهوم البنائي من نظرية الشخصية وحاجة الباحثين لطريقة تُحقّق صدق الأدوات المستخدمة في تطوير النظرية. ولا تركز أدلة المحتوى أو الأدلة المتعلقة بالمعيار بشكل مباشر على المفهوم البنائي الذي يجرى قياسه بالاختبار. والهدف من استخلاص دليل المفهوم البنائي هو تحديد المفهوم البنائي النفسي الذي يجري قياسه بالاختبار ومدى جودة قياسه.

تجمع دراسات المفاهيم البنائية أساليب منطقية وتجريبية. وأحد مظاهر الأسلوب المنطقي هو السؤال حول ما إذا كانت العناصر التي يقيسها الاختبار هي العناصر التي تكوّن المفهوم البنائي. مثلاً، حين ابتكر دول (Doll, 1935) مقياس فاينلاند vinelandللنضج الاجتماعي فإنه عرف المفهوم البنائي "النضج الاجتماعي" بكونه مجموعة

عناصر تبادلية الترابط هي مساعدة الذات، والتوجيه الذاتي، والتحرك والمهنة، والاتصال، والعلاقات الاجتماعية. إن من يراجعون الطبعة المنقحة الأولى للاختبار في كتاب Buros بعنوان (الكتاب السنوي للقياسات العقلية، 1949) يميلون إلى الاتفاق بأن هذه العناصر هي مظاهر للمفهوم البنائي يجب دمجها في اختبار النضج الاجتماعي. ويكون هناك أحياناً عدم اتفاق حول ماهية عناصر المفهوم البنائي. فمثلاً، إذا ظن أحدهم أن المفهوم البنائي "الذكاء" هو في الأساس مجموعة مهارات تمكّن الفرد من مواكبة بيئية أكاديمية، فإنه سوف يتوقع إن مثل هذه المهارات يمكن قياسها في اختبار ذكاء. ولو عرّف أحدهم "الذكاء" كمجموعة مهارات لا ترتبط بالبيئات المدرسية أكثر من بيئات أخرى فإنه سوف لن يرغب بدمج المهارات المدرسية الخاصة، في الاختبار.

وثمة مظهر آخر للأسلوب المنطقي ويكون بالمعاينة المدققة للبنود لتحديد ما إذا كانت تبدو مناسبة لتقييم عناصر المفهوم البنائي. ففي معيار فاينلاند، مثلاً، يُسأل والد طفل بعمر السادسة ما إذا كان طفله يستخدم الزلاجات والزحافة والعربة (المهنة) ويذهب إلى النوم دون مساعدة (المساعدة الذاتية) ويطبع كلمات بسيطة (الاتصال) ويلعب ألعاباً بسيطة (حركة) ويكون موضع ثقة عندما يكون لديه نقود (توجيه ذاتي). وتبدو هذه الأسئلة مناسبة لقياس عناصر النضج الاجتماعي. فلو تضمن الاختبار الأصلي أسئلة تتعلق بتفضيل الطفل لبعض الأغذية أو ما إذا كان يستخدم يده اليمنى لأمكن استبعاد مثل هذه البنود لأنها لا ترتبط بصورة مباشرة بالعناصر الخاصة بالمفهوم البنائي.

وتجمع البيانات التجريبية كدليل. فداخلياً، لا بد أن تكون العلاقات ضمن الاختبار كما يتنبأ بها المفهوم البنائي، وخارجياً، لا بد أن تكون العلاقات بين درجات الاختبار والملاحظات الأخرى متفقة مع المفهوم البنائي. لقد كان (دول) قادراً على إيضاح إن درجات فاينلاند للمهنة، والمساعدة الذاتية، وما إلى ذلك كانت مترابطة إيجابياً مع بعضها. وقد قدمت هذه الملاحظات الدعم الداخلي للنظرية التي تنص على أن مفهوم النضج الاجتماعي يتكون من عناصر داخلية مترابطة بالتبادل وقدمت الدليل بأن مقياس فاينلاند كان ناجحاً في قياس هذه العناصر المترابطة تبادلياً.

فإذا كانت علاقات العناصر في اختبار معين هي غير ما تنبأ بها المفهوم البنائي، فعندئذ إما أن يكون المفهوم البنائي، ذاته، غير مناسب أو إن الاختبار أخفق في قياس العناصر المتضمنة بالمفهوم.

مثلاً، قد ينوي أحدهم قياس المفهوم البنائي "الغريزة الاجتماعية البيولوجية" أولاً بافتراض أن المفهوم مكون من عناصر مترابطة بالتبادل: 1) إرادة البقاء، 2) إرادة الإنجاب، 3) رغبة اختيار القرين السليم، 4) رغبة التضحية بغية تعزيز البقاء فقط لأطفاله وأقرب أقربائه. فإذا وجد المرء – عند بناء وإجراء الاختبار – إن هذه العناصر غير مترابطة بشكل

إيجابي، فسوف يستنتج إن المقياس يفتقر إلى صدق المفهوم البنائي، ولذا ينبغي مراجعة الاختبار أو المفهوم البنائي ذاته.

وينبغي أن تكون درجات الاختبار مترابطة مع مقاييس خارجية بطريقة تتفق مع المفهوم. وقد وضح (دول) وآخرون أن الدرجات على مقياس فاينلاند تترابط فعلاً مع العمر الزمني، والعمر الذهني، ومع التقييمات المستقلة للنضج الاجتماعي. وعليه يمكن القول أن الدرجات على مقياس فاينلاند توضح العلاقات مع المقاييس الخارجية التي ينبغي توقعها في مقياس للنضج الاجتماعي الذي يتمتع بالصدق لقياس المفهوم البنائي.

ولابد أن يكون قياس مفهوم بنائي معين، مستقلاً قدر الإمكان عن قياسات المفاهيم البنائية الأخرى. فمثلاً، إذا طورنا اختباراً مصمماً لقياس مهارات حل المسائل الحسابية ووجدنا أن درجات هذا الاختبار مرتبطة جداً مع درجات اختبارات القراءة، فإننا سوف نستنتج بأننا طورنا اختباراً آخر للقراءة بدلاً من اختبار حل المسائل الحسابية ذاته. وربما يتعذر تطوير اختبار لحل المسائل الحسابية غير المترابط تماماً مع القراءة. وعلى أية حال، إذا كان لدينا اختباران متنافسان لحل المسائل الحسابية ودرسنا ارتباطهما مع اختبار للحساب، (r =0.7) إلا إن أحدهما مترابط مع اختبار للقراءة (r =0.8) والآخر مترابط مع اختبار القراءة (r =0.6) فأننا سوف نحكم على الأخير بأن له صدق أكبر في قياس حل المسائل الحسابية لأنه أكثر استقلالاً عن القراءة.

الطرق المستخدمة في استخلاص الدليل المتعلق بالمفهوم البنائي

Methods Used In Gathering Construct-Related Evidence

ليس هناك طريقة واحدة مستخدمة لاستخلاص الدليل المتعلق بالمفهوم البنائي لاختبار ما. فأي دليل له تأثير على معنى أو فائدة الدرجات، يكون مناسباً. ونعرض فيما يأتي بعض الأساليب المألوفة المستخدمة في استخلاص الأدلة المتعلقة بالمفهوم البنائي.

العلاقة مع مقاييس أخرى: ناقش (Messick, 1989) استخدام تقارب المؤشرات للمفهوم البنائي وكذلك قابلية التمييز عن مفاهيم بنائية أخرى. و "التقارب" يعني أن المقياس مترابط مع مقاييس أخرى يفترض أنها مؤشرات صادقة لذات المفهوم البنائي. فالمرء يبحث عن تقارب مؤشرات للمفهوم عن طريق السعي وراء مقاييس أخرى، ينبغي أن يترابط المفهوم نظرياً بها، ويبين بعدئذ كيفية ترابطها. فلابد لاختبار الاستدلال الرياضي أن يترابط مع الدرجات في مادة الرياضيات. فإذا ارتبط الاختبار بشكل كبير مع درجات الرياضيات فإن هذا دليل على التقارب.

ومع ذلك، فإن تقارب المؤشرات ليس بالدليل الكافي. فقد أشار ميسك إلى الحاجة

لدليل يمكنه تمييز المفهوم البنائي تجريبياً عن المفاهيم الأخرى. ولأجل إقامة قابلية التمييز يبحث المرء عن دليل يبين أن المفهوم البنائي لا يترابط جوهرياً مع أدوات معروفة لقياس مفاهيم بنائية أخرى مختلفة، أي أن المرء يحدد المقاييس التي لا ينبغي أن يترابط المفهوم البنائي معها بشكل جوهري. فلابد لاختبار الاستدلال الرياضي أن يتمتع بترابط متدن مع اختبار القراءة لأن القراءة، هي متغير لا علاقة له في اختبار الاستدلال الرياضي. فإذا وجد ترابط متدن بين اختبار الرياضيات واختبار القراءة فإن ذلك سيكون دليلاً على قابلية التمييز. وبالطبع، فإن ارتباط متدنيا أو صفراً مع أي مفهوم بنائي لن يكون دليلاً مناسباً. فلابد للمفهوم البنائي المستخدم أن يمثل على الأقل بعض المظاهر المحتملة للمفهوم البنائي قيد البحث. وبينما قد يكون من المعقول اختيار القراءة كمقياس لا ينبغي أن يترابط معه اختبار الاستدلال الرياضي بشكل جوهري فإنه لا جدوى من ربط درجات الرياضيات مع رمي الكرة لتقديم دليل متعلق بالمفهوم البنائي. وقد يساعد التفكير بمقياس للمفهوم البنائي المعني كممثل لنقطة على فترة متصلة للمفهوم المعني ومفاهيم بنائية للتقارب والتمييز كنقاط أخرى على الفترة المتصلة ذاتها. مثلا، يتوقع من مقياس الاختلاط الاجتماعي أن يترابط بشكل إيجابي مع مقياس الانبساط النفسي (التقارب)، وبشكل سلبي مع مقياس الانطواء (التمييز). ولمناقشة أكثر اكتمالا لدور التقارب والتمييز كأدلة على المفهوم البنائي المقاس، ندعو القارئ إلى الرجوع إلى مقالة (Compbell & Fiske, 1959) القديمة.

وثمة مظهر آخر لطريقة الترابط لجمع الأدلة هي التحليل العاملي. والتحليل العاملي هو طريقة إحصائية لدراسة الترابطات المتبادلة بين مجموعة من درجات الاختبارات بغية تحديد عدد العوامل (المفاهيم البنائية) المطلوبة لتفسير هذه الترابطات المتبادلة. وتقدم الطريقة كذلك معلومات عن العوامل التي تحدد الأداء في كل اختبار وكذلك النسبة المئوية للتباين في درجات الاختبار التي تفسرها العوامل. فالمرء يبدأ بعدد كبير من المقاييس المختلفة وبفحص الترابطات بينها وإيجاد تلك المقاييس التي تتماشى (تترابط) مع بعضها، قد يقلص المرء من العدد الكبير للدرجات إلى عدد أصغر من العوامل التي يجرى قياسها وتعد أساساً لها. ولا توضح الترابطات الاختبارات التي تقيس العامل ذاته، فحسب، بل كذلك مدى قياس ذلك العامل. وبفحص محتوى الاختبارات التي تترابط مع العامل نفسه يمكن للمرء أن يستدل طبيعة المفهوم البنائي المقاس.

دراسات تجريبية: قد يفترض أن درجات الاختبار تتغير عندما يتم إدخال أنواع معينة من المعالجات التجريبية. فمثلاً، في إضفاء الصدق على مقياس القلق، قد يفترض المرء أن الدرجات على المقياس سوف تتغير عندما يوضع الأشخاص في موقف مثير للقلق. وإذا

جرى تفعيل القلق في تجربة ضابطة(*) وتغيرت الدرجات الناتجة في مقياس القلق بالطريقة المتنبأ بها، فسوف يكون لدينا دليل ما بأن المقياس يقيس القلق، فعلاً.

مقارنة درجات مجموعات محددة: يمكن لنا استخدام مجموعات معروفة مسبقاً باختلافها والافتراض بأن الدرجات حسب الأداة المعنية سوف تميز مجموعة عن أخرى. وقد يتوقع المرء أن الدرجات في اختبار الاستعداد الموسيقي سوف تميز بين الطلبة المسجلين حالياً في مدرسة الموسيقى ومجموعة ما من طلبة الكلية. وبصورة مماثلة إذا أمكن التمييز بين الميكانيكيين وغير الميكانيكيين على أساس درجاهم في اختبار الاستعداد الميكانيكي فإن ذلك سوف يقدم دعماً لصدق الاختبار باعتباره مقياساً للاستعداد الميكانيكي. وإذا قاس استبيان التوافق النفسي، فمن المتوقع لدرجات الاستبيان أن تميز بين المجموعات المحددة مسبقاً على أها متوافقة (سويّة) وتلك المحددة مسبقاً على أها عصابية. وقد تكون المجموعات المختلفة المستخدمة للمقارنة في دراسات الصدق هي مجموعات (تعود للعمر أو الجنس) أو لمقادير مختلفة من التدريب في مجال معين له علاقة بالمفهوم البنائي ومجموعات معروفة بكوها سوية وأخرى معروفة بأها سيئة التوافق وما إلى ذلك. فإن تم تأكيد الفروقات المتنبأ بها في درجات الاختبار فسيكون لدينا دعم لصدق الاختبار باعتباره مقياساً للمفهوم البنائي المعني.

التحليل الداخلي للاختبار: تفحص طريقة التحليل الداخلي للاختبار، ذاته، وتجمع معلومات عن محتوى الاختبار والعمليات المستخدمة في الاستجابة على بنود الاختبار والترابطات بين بنود الاختبار. وقد تعرض البيانات من الدراسات المتعلقة بالمحتوى معلومات مناسبة حول المفهوم البنائي المقاس بالاختبار. فعند تحديد مجتمع سلوكي يجرى الاختبار على عينة منه، قد يتوفر بعض الفهم لطبيعة المفهوم البنائي الذي يجرى قياسه بالاختبار. فمثلاً إذا كان لابد من تحديد مجتمع سلوكي لاختبار استدلالي عن طريق وصف القدرات التي يجرى تكوين عينتها عن طريق الاختبار (كالقدرة على فهم المتشابهات الكمية واللفظية) فإننا تحصل على بعض من الفهم حول صدق الاختبار.

وقد يبحث المرء في العمليات العقلية والمهارات التي يستخدمها الأشخاص لدى الاستجابة لبنود الاختبار. فمثلاً، قد يطلب من الطلبة "التفكير بصوت عال" حينما يعملون خلال اختبار التفكير اللفظي. وقد يكشف مثل هذا الإجراء بأن الاختبار يقيس قدرة الاستدلال هذه كما يزعم، أو قد يكشف أن عوامل أخرى كالمفردات أو استيعاب القراءة يتم قياسها.

(*) التجربة الضابطة (Controlled Experiment): تجربة يجرى بموجبها ضبط المتغيرات المستقلة وتغيير المتغيرات التابعة أثناء سير التجربة (قاموس التربية - الخولي) - (المراجع)

وينبغي البحث في تجانس محتوى الاختبار بغية التأكد ما إذا كان الاختبار يقيس سمة أو صفة واحدة. إن قياسات التوافق / الثبات الداخلي كمعامل ارتباط ألفا أو معامل ارتباط كودر - ريتشاردسون تقدم دليل التجانس. وستناقش هذه المقاييس في جزء تال من هذا الفصل. وتقدم مقاييس التجانس دليلاً متعلقاً بالمفهوم البنائي إذ أنها تساعد في وصف المدى الذي تقاس عنده سمة أو مفهوم بنائي واحد. ومع ذلك، فإن بيانات الاختبار الداخلي غير كافية أبداً، لإضفاء الصدق على الاختبار. فنحن نحتاج إلى بيانات خارجية بغية تحديد ما يقيسه الاختبار.

إسهام الدراسات المتعلقة بالمفهوم البنائي

Contribution of Construct-Related Students

يعد هذا النوع من الدراسات الأكثر شمولاً، لأنها تصنف صلة المحتوى وتمثيله وكذلك الدليل المتعلق بالمعيار. وتعد الطريقة المتعلقة بالمفهوم البنائي لصدق الاختبار مهمة لأنها تركز الاهتمام على دور النظرية في بناء الاختبار والحاجة إلى صياغة فرضيات يمكن بحثها كجزء من عملية تحقيق الصدق.

وحتى إذا لم يتضمن الأمر نظرية محددة بشكل، مباشر فإن هذه الطريقة تؤكد على الحاجة إلى ربط المفهوم البنائي المعني باختبار معين في إطار مفاهيم يحدد معنى المفهوم البنائي ويميزه عن المفاهيم البنائية الأخرى، ويوضح كيف ينبغي لمقاييس المفهوم البنائي الارتباط مع المتغيرات الأخرى. وهكذا يمكن للباحث أن يجمع البيانات من مختلف المصادر لتقديم الإسناد لصدق اختبار المفهوم البنائي.

مقارنة مصادر الدليل من أجل الصدق

Comparison of The Sources of Evidence for Validity

إن تنوع أدلة الصدق تكمل بعضها ببعض وتتعلق جميعها جوهرياً بالتفسيرات المناسبة لمعنى الدرجات وكيفية استخدامها. ولهذا السبب يأتي إدراك الصدق كمفهوم موحد.

أفترض إن مدرساً أراد بناء اختبار قراءة يستخدم مع الصف السادس. فما نوع الدليل الذي ينبغي على المدرس استخلاصه لدعم صدق استخدامات الاختبار؟ فإذا كان القصد من هذا الاختبار، التوصل لإجراء استنتاجات عن تحصيل القراءة، فإن المدرس سيتخذ أول قرار حول مجتمع المحتوى الذي ستؤخذ عينة منه. وبشكل خاص، يتم تحديد الكتب المقررة، ومواد القراءة الخارجية، والتمارين الصفية وما إلى ذلك. ويتم اختيار عينات من هذا المجتمع بطريقة تضمن تكوين عينة ممثلة للمحتوى كله ولأهداف المقرر. كما سيطلب من مدرسين آخرين إجراء الحكم الخاص بكفاية المحتوى للهدف المقصود.

ويستخلص الدليل حول العلاقة بين درجات الطلبة في الاختبار وأدائهم حسب معيار مناسب. ويمكن إجراء دراسة متابعة لتحديد ما إذا كان هناك ارتباط بين درجات الاختبار هذه وتحصيل القرائي في الصف السابع. ويقدم الارتباط العالي دليلاً على أن لاختبار القراءة صدقاً في التنبؤ بتحصيل القراءة في الصف السابع. كما يمكن الحصول على دليل إضافي عن طريق ترابط درجات الاختبار مع درجات الطلبة في درس القراءة أو مع درجات اختبار قراءة مقنن صادق.

ولغرض استخلاص دليل الصدق المتعلق بالمفهوم البنائي، يحاول المدرس تحديد ما إذا كانت الفرضيات الخاصة بطبيعة كفاءة القراءة مسندة بأداء الطلبة في الاختبار. وسوف تتم دراسة الفروقات في أداء ذوي الدرجات العالية مقابل ذوي الدرجات المتدنية، كما ستجرى محاولة لتحديد ما إذا كانت لأداء الاختبار صلة بنظرية لتدريس القراءة وما إلى ذلك. وهناك طريقة أخرى تنطوي على دراسة تقارب درجات القراءة مع متغيرات أخرى. إن قياس الثبات / الاتساق سوف يوضح ما إذا كان الاختبار يقيس مفهوماً بنائياً واحداً. إن خلاصة لجميع أنواع الأدلة أعلاه سيوضح مدى ملاءمة وفائدة الاستنتاجات من درجات اختبار القراءة، ومن ثم صدق الاختبار، ويلخص الجدول (8.2) دليل الصدق المتعلق بالأغراض المختلفة للاختبارات.

الجدول 8.2: صدق اختبار قراءة لأغراض مختلفة

نوع الدليل	مثال	أسئلة تطرح
المحتوى	اختبار تحصيل في قراءة للصف السادس	ما مدى قيام الاختبار باختبار ما جرى تعليمه؟
المتعلق بالمعيار: دراسة تنبؤية	اختبار استعداد للتنبؤ بالأداء في قراءة الصف السابع	ما مدى قيام الاختبار بالتنبؤ حول تحصيل قراءة الصف السابع؟
المتعلق بالمعيار: دراسة متزامنة	اختبار لتحديد أطفال ما قبل المدرسة المعرضين للخطر	ما مدى توافق الاختبار مع ملاحظات أخرى للأطفال؟
المفهوم البنائي	اختبار لقياس استيعاب القراءة	ما مدى قياس اختبار المفهوم البنائي "استيعاب القراءة"؟ هل تسند البيانات، والفرضيات حول استيعاب القراءة؟

تطبيق مفهوم الصدق
Application of The Validity Concept

مع أننا نعرّف "الصدق" بطريقة عامة بكونه "المدى الذي يقيس عنده اختبار معين ما يراد قياسه"، إلا أن الصدق ليس خاصة عامة يحوزها الاختبار. إن الصدق خاص بعمل معين يريد أحدهم من الاختبار أن يقوم به. وينبغي أن يقوم بناء الاختبار أو المقياس على غرض واحد في الذهن. فأجراء يراد منه أغراض متعددة سوف لن ينجز أياً منها بشكل جيد. يجب على المرء أن يعرف الغرض من الاختبار والإطار والمجتمع الذي يستخدم فيه بغية تقييم صدق الاختبار لتلك الظروف الخاصة.

الثبـــات / الاستقرار RELIABILITY

لقد ذكرنا في بداية الفصل، بأن الثبات لأداة قياس يشير إلى درجة التوافق / الاستقرار مع ما تقيسه، مهما كان ما تقيس.

وتعد هذه الصفة أساسية في أي نوع من القياس. فدائرة بريد ستتخذ إجراءاً في الحال لإصلاح مقياس إذا وجدت أن هذا المقياس يقلل أو يزيد أحياناً من وزن الطرود. كذلك، فإن علماء النفس والتربويين معنيون بالمثل، بتوافق / ثبات أدوات قياسهم حين يحاولون قياس صفات معقدة مثل الاستعداد الدراسي والتحصيل والدافعية والقلق وما شابه. فهي لن تعتبر اختبار الاستعداد الدراسي مجدياً إذا تمخض عن نتائج مختلفة بشكل واضح في كل مرة يستخدم فيها للموضوع ذاته. فيجب على الناس الذين يستخدمون مثل أدوات القياس هذه، أن يحددوا ويستفيدوا من الأساليب التي تساعدهم في تحديد مدى توافق / ثبات أدوات القياس وموثوقيتها.

نظرية الثبات Theory of Reliability

كطريقة للتمييز بين مفهوم الثبات ومفهوم الصدق، من المفيد تحديد الأخطاء العشوائية للقياس والأخطاء المنتظمة للقياس. فالخطأ العشوائي يشير إلى الخطأ الناجم عن الصدفة المحضة. وأخطاء القياس العشوائي قد تضخم أو تخفض درجة الفرد بطريقة لا يمكن التنبؤ بها. فمثلاً، أحد العناصر في اختبار اللياقة البدنية لتلاميذ الابتدائية هي رمية كرة البيسبول. فيوصي الأفراد برمي الكرة إلى أبعد ما يستطيعون ثم تقاس مسافة الرمية، ومع إن موضوع الاختبار هو تسجيل درجة نموذجية لأداء الفرد، وبالتأكيد إذا كان لدينا فرد واحد يرمي الكرة في عدة مناسبات، سنجد بأن الطفل لا يرميها إلى ذات المسافة في كل مرة.

إفترض إن كل طالب قام برمية على مدى يومين متعاقبين. فإذا قارنا الدرجتين

(مسافتي الرمي) لكل طالب سنجد، تقريباً، أنها ليست نفسها بالضبط. إن أغلب الفروقات ستكون صغيرة لكن بعضها سيكون كبيراً بشكل معتدل والقليل منها سيكون كبيراً جداً. إن النتائج غير متوافقة من رمية في يوم لرمية في يوم آخر. لذا فإن رمية معينة لا يعوّل عليها ثابتة تماماً كقياس لقدرة الرمي عند الطالب.

هناك ثلاثة أنواع للصدفة، أو العشوائية، التي تؤدي إلى عدم التوافق بين الدرجات المتحققة في اليومين.

1- "قد يتغير الطالب فعلاً"، من يوم لآخر. ففي يوم معين قد يشعر أنه أفضل من اليوم الآخر. وفي يوم معين قد يكون الطالب أكثر دافعية وأقل تعباً. وربما يبدأ والد الطفل الذي يسمع عن المهمة، بتدريبه على رمي كرة البيسبول.

2- "قد يتغير العمل نفسه" في القياسين. مثلاً، قد تكون الكرة المستخدمة في يوم صلبة بينما قد تكون رطبة مبتلة بالماء في اليوم الثاني. وربما يسمح الفاحص في يوم معين للطلبة بالركض حتى خط الرمي بينما في اليوم التالي يسمح لهم فقط بخطوتين. إن هذه التغيرات قد تساعد بعض الطلبة أكثر من غيرهم.

3- إن "العينة المحدودة" للسلوك تؤدي إلى درجة غير مستقرة. فالعينة الصغيرة للسلوك عرضة للعديد من تأثيرات الصدفة. فربما تكون هناك عصفة ريح عند رمي الكرة. وربما يفقد الطالب توازنه عند الشروع برمي الكرة أو ربما انزلقت أصابعه أثناء مسك الكرة.

فالثبات معني بأثر الأخطاء العشوائية للقياس على توافق / استقرار الدرجات.

ومن ناحية أخرى، فإن بعض الأخطاء الخاصة بالقياس قابلة للتنبؤ أو منتظمة. وباستخدام مثال رمية كرة بيسبول، تصور موقفاً تعطى فيه تعليمات الرمي بالإنكليزية حيث لا يفهم جميع الأفراد الإنكليزية. فدرجات غير الناطقين بالإنكليزية يمكن أن تكون منخفضة بشكل منتظم لأن الأفراد لا يفهمون ما يتوقع منهم القيام به. فمثل هذه الأخطاء المنتظمة للقياسات هي مشكلة صدق. ويتدنى صدق الاختبار كلما تغيرت الدرجات بصورة منتظمة بتأثير شيء ما، خلافاً لما نحاول قياسه. وفي هذا المثال، فإننا لا نقيس قدرة رمي كرة البيسبول فحسب، بل كذلك فهم الإنكليزية ولو جزئياً.

ولاتخاذ قرار حول ما إذا كنا نتعامل مع الثبات أو الصدق، علينا أن نقرر ما إذا كنا سنأخذ الأخطاء العشوائية أو الأخطاء المنتظمة بنظر الاعتبار. فإذا أعطي لصف اختبار لرمي الكرة واستخدمت كرتان، واحدة صلبة والأخرى مبتلة بالماء واعتمد الاختبار مسألة الصدفة عمن يحصل على أية كرة منهما، فإن التباين الناتج عن الكرة المستخدمة هو مشكلة ثبات. فالتباين الذي تتسبب به الكرة يمثل خطأ عشوائياً يؤثر على توافق /

ثبات القياسات. ولو طلب من تلاميذ الصف أن يخضعوا للاختيار حسب سياق أبجدي وكان اليوم ماطرا والكرة المستخدمة تزداد بللاً مع كل رمية تالية، فإن التباين بسبب الرطوبة المتزايدة للكرة سيكون مشكلة صدق. فالدرجات، في هذه الحالة، تكون مرتفعة لمن هم قرب بداية الأبجدية، ومنخفضة بالنسبة لمن هم قرب نهايتها. إن صدق درجات رمي الكرة تقل لأن الدرجات لا تعكس فقط شدة رمي الكرة بل السياق الأبجدي كذلك. وهذا مثال على الخطأ المنتظم الذي يؤثر على صدق القياس.

فالثبات يتعلق بمدى ثبات / استقرار قياسنا لكل ما نقيسه. وهو لا يتعلق بما إذا كنا نقيس ما ننوي قياسه – فهذه هي مسالة الصدق. فمن الممكن بالنسبة لأداة القياس أن تكون موثوقة بثباتها دون أن تكون صادقة. وعلى أية حال، لا يمكن لها أن تكون صادقة ما لم تكن تتمتع بالثبات أولاً. فمثلاً، يمكن لأحدهم أن يقرر القيام بقياس الذكاء عن طريق تحديد محيط الرأس. وقد تكون القياسات متوافقة من وقت إلى آخر (ثابتة) لكن هذه الطريقة لا تعتبر قياساً صادقاً للذكاء، إذ أن محيط الرأس لا يترابط مع أية معايير أخرى للذكاء ولا يخضع للتنبؤ في أية نظرية للذكاء.

يتأثر الثبات بالأخطاء العشوائية والتي هي أية عوامل تنجم عنها تباينات بين الدرجات في محاولات متكررة لأداة قياس واحدة أو من مجموعة بنود إلى مجموعة مكافئة لها.

وتبرز الأخطاء العشوائية من عدد من المصادر. وقد تكون الأخطاء ملازمة للوسيلة ذاتها. فمثلاً، إذا كان الاختبار قصيراً فإن الأفراد الذين يتصادف أنهم يعرفون الإجابات القليلة سيحصلون على درجات أعلى مما يستحقونه، بيد أن أولئك الذين لا يعرفون تلك الإجابات القليلة سينالون درجات أقل مما يستحقونه. فمثلاً لو أعطي اختبار لتقييم مدى معرفة الطلبة بعواصم الخمسين ولاية بطرح خمسة أسئلة فقط، فإن من المحتمل أن الطالب الذي يعرف عشرة عواصم فقط قد يعطي خمسة أجوبة صحيحة، بينما الطالب الذي يعرف (40) قد لا يعطي أي جواب صحيح. ففي الاختبار القصير يكون الحظ عاملاً أكثر مما هو عليه في الاختبار الطويل. وإذا كان الاختبار سهلاً بحيث يعرف كل طالب أغلب الأجوبة فإن الدرجات النسبية للطلبة تعتمد مرة أخرى فقط على أسئلة قليلة، ويكون الحظ عاملاً كبيراً. وإذا كانت الأسئلة غامضة فإن الطلبة "المحظوظين" سوف يجيبون بالطريقة التي أرادها الفاحص بينما الطلبة "غير المحظوظين" سيجيبون بطريقة صحيحة مماثلة إلا إن إجاباتهم سوف تسجل على أنها غير صحيحة. إن إجراء التصحيح يؤثر كذلك على الثبات. أما إجراءات التصحيح الدقيق فإنها تعزز الثبات بينما تضعفه إجراءات التصحيح الغامضة.

قد تكون الأخطاء ملازمة في تطبيق الأداة. فقد يبتعد الشخص عديم التجربة عن

الإجراءات المقننة في تطبيق الاختبار أو تصحيحه. وقد تؤثر ظروف الاختبار كالضوء والحرارة والتهوية على الأداء. وقد تكون تعليمات إجراء الاختبار غامضة.

وهناك أيضاً خطأ التلميذ – أي التذبذب في الدافعية، والاهتمام، والتعب، والظرف البدني، والقلق، وعوامل عقلية وانفعالية أخرى تؤثر في نتائج الاختبار. فالتلميذ الذي يكسر نبلة القلم في اختبار موقوت سوف يزيد من عنصر الخطأ في النتائج.

معادلات للثبات Equations For Reliability

من المقبول عموماً أن تحتوي جميع قياسات الصفات الإنسانية على بعض الخطأ. فإجراءات الثبات تتعلق بتحديد درجة عدم التوافق في الدرجات بسبب الخطأ العشوائي.

حين يجري أحدهم اختباراً لطالب فإنه يضمن درجة يمكن أن تدعى الدرجة الملاحظة / الخام. وإن تعين على أحدهم اختبار هذا الطالب في مناسبة أخرى بالأداة ذاتها فإنه، على الأرجح، لن يحصل بالضبط على الدرجة الملاحظة ذاتها. والدرجة الملاحظة تحتوي على خطأ قياس. وعليه، يستنتج المرء إن كل درجة اختبار تتكون من عنصرين: الدرجة الحقيقية مضافاً لها بعض الخطأ في القياس. وكما أشرنا أعلاه، فإن عنصر الخطأ قد يعزى مجموعة أو عدد من العوامل مرتبطة بالتباينات في الشخص من وقت لآخر أو بإجراء الاختبار على ذلك الشخص.

ويعبر عن ثبات الاختبار رياضياً، كأفضل تقدير لنسبة التباين الحقيقي من التباين الكلي لدرجات الاختبار. وكما أوضحنا في الفصل الخامس فإن التباين هو مؤشر انتشار مجموعة من الدرجات. فإذا أجرينا اختباراً لمجموعة من الطلبة فإن بعض انتشار (تباين) درجات الطلبة يعزى إلى فروقات حقيقية بين المجموعة وبعض الانتشار (التباين) يعزى إلى أخطاء القياس.

إن فكرة مكوّن / عنصر الخطأ والمكوّن / العنصر الحقيقي في درجة اختبار واحدة قد يتم تمثيلها بشكل رياضي في المعادلة 8.1:

$$X = T + E \qquad (8.1)$$

حيث:

X = الدرجة الملاحظة.
T = عنصر الدرجة الحقيقية.
E = عنصر خطأ القياس.

قد يعرّف عنصر الدرجة الحقيقية على أنها الدرجة التي ينالها شخص معين تحت ظروف تستخدم فيها أداة قياس متقنة، بينما يمكن لعنصر خطأ القياس أن يكون موجباً أو

سالباً. فإذا كان موجباً فإن الدرجة الحقيقية للشخص ستكون عالية التقدير في درجة الملاحظة وإذا كان سالباً ستكون متدنية التقدير. وبسبب افتراض أن خطأ القياس يمكن أن يكون موجباً وسالباً فإنه يمكن الاستنتاج بأن كلا من مجموع الأخطاء ومتوسط الخطأ سيكون صفراً فيما إذا طبقت أداة القياس ذاتها أو ما يكافؤها على شخص لعدد غير منته من المرات. وتحت هذه الظروف، يعرّف المكون الحقيقي على أنه درجة الوسط الحسابي للشخص في عدد غير منته من القياسات. والدرجة الحقيقية هي مفهوم نظري طالما أن العدد غير المنتهي لإجراء الاختبار على نفس الشخص غير عملي.

وفي الوضع الاعتيادي للبحوث، يكون لدى الباحث قياس واحد لكل مجموعة أشخاص. وبعبارة أخرى يكون لدى الباحث مجموعة واحدة من درجات الاختبار يأخذها بنظر الاعتبار. ولكل درجة مرصودة / ملاحظة عنصر درجة حقيقية وعنصر درجة خطأ. لقد تبين رياضياً إن تباين الدرجات المرصودة لمجموعة كبيرة من الأشخاص (σ_x^2) مساو لتباين درجاتهم الحقيقية (σ_t^2) مضافاً له التباين في أخطاء القياس (σ_e^2) أو

$$\sigma_x^2 = \sigma_t^2 + \sigma_e^2 \qquad (8.2)$$

وقد يعرف الثبات نظرياً بأنه النسبة بين الدرجات الحقيقية وتباين الدرجات المرصودة في مجموعة درجات كما تعبر عنه المعادلة التالية:

$$r_{xx} = \frac{\sigma_t^2}{\sigma_x^2} \qquad (8.3)$$

حيث:

r_{xx} = ثبات الاختبار.

σ_t^2 = تباين الدرجات الحقيقية.

σ_x^2 = تباين الدرجات المرصودة.

فالثبات هو الجزء من التباين في الدرجة المرصودة الخالية من الخطأ، ويمكن التعبير عن هذه الفكرة بالمعادلة التالية المستمدة من المعادلتين (8.2) و (8.3):

$$r_{xx} = 1 - \frac{\sigma_e^2}{\sigma_x^2} \qquad (8.4)$$

فيمكن لمعامل الثبات (r_{xx}) أن يتراوح من (1) حين لا يكون هناك خطأ في القياس إلى صفر، حين يكون القياس كله خاطئ. (حين لا يكون هناك خطأ في القياس تكون (σ_e^2) في المعادلة السابقة صفراً و $r_{xx} = 1$. وإن كان القياس جميعه خطأ، فإن $\sigma_e^2 = \sigma_x^2$ و $r_{xx} = 0$. ويتضح مدى الخطأ بمدى ابتعاد معامل الثبات عن (1). فالعامل (80.) في اختبار معين مثلاً يوضح أن أفضل تقدير هو أن (80%) من التباين الموجود في الدرجات هو تباين

حقيقي وأن (20%) خاطئ. وهكذا كلما كان الخطأ كبيراً تدنى معامل الثبات عن (1) بثبات أدنى. وبالعكس كلما كان معامل الثبات قريباً من (1) كان للأداة خطأ قليل نسبياً وثبات عال.

طرق الثبات Approaches to Reliability

يعتبر الاختبار ثابتاً / مستقراً حسب المدى الذي تظل فيه الدرجات، التي ينالها الشخص، نفسها تقريباً في قياسات متكررة. وهناك طريقتان للتعبير عن ثبات مجموعة قياسات. توضح الطريقة الأولى مقدار التباين المتوقع داخل مجموعة قياسات متكررة لشخص واحد. فإذا كان من الممكن وزن / تقدير شخص على مقياس (200) لحصلنا على توزيع تكراري لدرجات تمثل وزنه / تقديره. وسيكون للتوزيع التكراري قيمة متوسطة يمكن اعتبارها الوزن / التقدير "الحقيقي". وسيكون له أيضاً انحراف معياري يوضح الانتشار. ويدعى هذا الانحراف المعياري بخطأ القياس المعياري لأنه الانحراف المعياري "لأخطاء" قياس الوزن/ التقدير لشخص واحد. ولا نجري في الغالب، مع البيانات النفسية والتربوية، قياسات متكررة على الفرد. فالوقت لا يسمح لمثل هذا التكرار، وإضافة إلى ذلك، فإن الممارسة وأثار التعب المرتبطة بقياس متكرر، سيكون لها تأثير على الدرجات. وهكذا، فبدلاً من قياس شخص واحد عدة مرات، نقيس مجموعة كبيرة متنوعة في مناسبتين، وباستخدام زوج من القياسات لكل شخص يمكننا أن نقدر ما الذي سيكون عليه هذا الانتشار للشخص العادي لو أتيح لنا إجراء القياس مرة تلو الأخرى.

ويوضح ثبات القياس في الطريقة الثانية المدى الذي يحافظ عنده الفرد على الموضع النسبي ذاته في المجموعة. ومع اختبار يتمتع بالثبات / الاستقرار ينبغي أن يكون الشخص الذي ينال أعلى الدرجات في اختبار معين، اليوم، هو الشخص ذاته الذي سينال أعلى الدرجات في اليوم التالي إذا ما طبق عليه الاختبار ذاته. وسيحافظ كل شخص في المجموعة على الوضع النسبي ذاته تقريباً. وكلما تغير الأشخاص في موضعهم النسبي، كلما كان ثبات الاختبار متدنيا. وبوسعنا حساب معامل الارتباط بين إجراءين للاختبار ذاته لتحديد المدى الذي يحافظ منه الأشخاص على نفس الموضع النسبي. ويدعى معامل الارتباط هذا بمعامل الثبات (r_{xx}). فمعامل الثبات (1.00) يوضح أن الموضع النسبي لشخص في إجرائين يظل نفسه وإن الاختبار ثابت جداً.

وهكذا فتوافق / ثبات المقياس يتضح بخطئه المعياري للقياس أو بمعامل ثباته.

مؤشرات الثبات The Reliability Indexes

يمكن تقدير الثبات بربط الدرجات التي يحصل عليها الأشخاص أنفسهم في مناسبات

مختلفة أو مع مجموعات مختلفة من البنود المكافئة. وتتطلب هذه الإجراءات تطبيقين للاختبار. وتفحص إجراءات أخرى التوافق / الثبات الداخلي للاختبار وتتطلب تطبيقاً واحداً.

ثبات الاختبار – إعادة الاختبار(*) Test - Retest Reliability

إن الطريقة الواضحة لتقدير ثبات الاختبار تكون بإجرائه على المجموعة نفسها من الأشخاص في مناسبتين ثم دراسة ترابط الدرجات المزدوجة. ويدعى معامل الارتباط الناتج عن هذا الإجراء "معامل ثبات الاختبار – إعادة الاختبار". مثلاً، قد يعطى اختبار اللياقة البدنية إلى صف خلال أسبوع ثم يعطى الاختبار ذاته مرة أخرى في الأسبوع التالي. فإذا كان للاختبار ثبات جيد فإن الموضع النسبي للشخص في إجراء الاختبار الثاني سيكون قرب موضعه النسبي في إجراء الاختبار الأول. ويشار أحياناً إلى معامل ثبات إعادة الاختبار بوصفه "معامل الاستقرار"، وذلك لكونه مؤشراً لتوافق / ثبات درجات الشخص طوال وقت إضافي. فهو يبين ما إذا كان بوسعنا التعميم من درجة يحصل عليها شخص في مناسبة معينة إلى درجة سيحصل عليها ذلك الشخص إذا ما أعطي الاختبار في وقت مختلف.

يفترض معامل الاختبار – إعادة الاختبار أن الخاصة المقاسة بالاختبار مستقرة خلال الوقت، لذا فأي تغيير في الدرجات من وقت إلى آخر يعزى إلى خطأ عشوائي. وقد يعزى الخطأ إلى ظرف الأشخاص أنفسهم أو ظروف إجراء الاختبار. ويفترض معامل الاختبار – إعادة الاختبار أيضاً عدم تأثير المران أو الذاكرة. مثلاً، قد يتعلم الطلبة شيئاً من مجرد إجراء الاختبار، لذا سوف يستجيبون بشكل مختلف في إجراء الاختبار للمرة الثانية. إن تأثيرات المران / الممارسة هذه من الاختبار الأول لا يحتمل لها أن تكون متماثلة لدى جميع الطلبة، مما يقلل من تقدير الثبات. ولو كانت الفترة الزمنية قصيرة فقد يكون هناك تأثير الذاكرة إذ قد يجيبون على السؤال بذات الطريقة السابقة لمجرد أنهم يتذكرون الإجابة عليه بتلك الطريقة أول مرة. ويميل أثر الذاكرة إلى تضخيم تقدير الثبات، غير أنه يمكن السيطرة عليه، نوعاً ما، بزيادة الوقت بين الاختبار الأول وإعادة ذلك الاختبار. ومن ناحية أخرى، إذا كان الوقت بين الاختبارات طويلاً جداً، فإن التعلم الفارق قد يكون مشكلة – أي إن الطلبة سوف يتعلمون مقادير مختلفة خلال الفترة بما سيؤثر على معامل الثبات.

ونظراً للمشكلات التي نوقشت أعلاه، فإن إجراء الاختبار – إعادة الاختبار لن يكون مناسباً على الأغلب، بالنسبة للاختبارات في المجال المعرفي. فاستخدام هذا الإجراء في المدارس يتم حصره بشكل واسع، بمقاييس اللياقة البدنية والبراعة الرياضية.

(*) قد يطلق على مصطلح "الاختبار – إعادة الاختبار"، عبارة "إعادة الاختبار"، اختصاراً – (المراجع)

ثبات الأشكال / الصور المتكافئة Equivalent - Forms Reliability

يستخدم أسلوب الأشكال / الصور المتكافئة لتقدير الثبات الذي يدعى كذلك "أسلوب الصور المتبادلة" أو "أسلوب الصور المتوازية"، عندما يكون محتملاً أن يتذكر الأفراد استجاباتهم في بنود الاختبار. وبدلاً من ربط نتائج الإجراءين لذات الاختبار للمجموعة ذاتها، يمكن ربط نتائج صور متكافئة من الاختبار الذي يطبق على ذات الأشخاص. فإذا طبقت الصورتان، جوهريا، في ذات الوقت (بتعاقب مباشر) فإن معامل الثبات الناتج يدعى "معامل التكافؤ". ويعكس هذا الإجراء التباينات في الأداء من مجموعة محددة من البنود إلى أخرى، وهو يبين ما إذا كان بوسعنا تعميم درجة الطالب إلى ما قد يحصل عليه إذا ما أعطي صورة أخرى للاختبار ذاته. والسؤال هو ما المدى الذي يعتمد فيه أداء الطالب على المجموعة الخاصة من البنود المستخدمة في الاختبار؟

إذا تم اختبار الأشخاص بإحدى صور الاختبار في مناسبة واحدة، وبصورة مكافئة في مناسبة ثانية، وأجرى ارتباط بين درجات الصورتين فإن معامل الارتباط الحاصل يدعى "معامل الثبات والتكافؤ". ويعكس هذا المعامل مظهرين لثبات الاختبار: تباينات الأداء من وقت لآخر، وكذلك التباينات من صورة إلى أخرى. ويوضح معامل الثبات والتكافؤ المرتفع أن كلتي الصورتين تقيسان ذات الأداء وتقيسان الثبات خلال الفترة الزمنية. ويعد هذا الإجراء الأكثر دقة ومطلباً لتحديد ثبات الاختبار.

إن تصميم الصور المتبادلة لاختبار، التي تكون متكافئة حقا، هو المشكلة الكبرى مع هذا الأسلوب من تقدير الثبات. فإن لم يتحقق ذلك بشكل ناجح، فإن التباين في الدرجات من صورة إلى أخرى لا يمكن اعتباره تباين خطأ. فالصور المتكافئة لاختبار هي اختبارات يجرى بناؤها بشكل مستقل حيث لابد لها أن تلبي ذات المواصفات – أي لا بد أن يكون لها العدد ذاته من البنود، والشكل، والتعليمات، والحدود الزمنية، والتصميم، والمحتوى والمدى، ومستوى الصعوبة – إلا إن الأسئلة الفعلية ليست نفسها. ونموذجياً، لابد أن يكون لدى المرء أزواجاً من البنود المتكافئة ويخصص واحد من كل زوج لكل صورة. ففي اختبار جغرافية العالم مثلاً، قد يسأل الشكل أ: "في أية قارة يوجد نهر النيل؟" بينما قد يسأل الشكل ب: "في أي قارة يوجد نهر الأمازون؟" وقد يسأل الشكل أ، ما هي عاصمة إيطاليا؟ والشكل ب، ما هي عاصمة فرنسا؟ أما توزيعا درجات الاختبارين فينبغي لهما أن يكونا متكافئين.

ويوصى بأسلوب الصور المتكافئة عندما يرغب المرء في تجنب مشكلة أثر التذكر أو المران، وفي الحالات التي يكون فيها لدى المرء عدد كبير من بنود الاختبار يمكن فيها اختيار عينات متكافئة. ويعتبر، عموما، أن إجراء الصور المتكافئة يؤمن أفضل تقدير لثبات المقاييس الأكاديمية والنفسية.

مقاييس التوافق الداخلي للثبات

Internal - Consistency Measures of Reliability

وهناك أساليب ثبات أخرى، يجرى تصميمها لتقرير ما إذا كانت كل البنود في الاختبار تقيس الشيء ذاته. وتدعى هذه "أساليب التوافق / الثبات الداخلي" وتتطلب تطبيقاً واحداً لصورة واحدة من الاختبار.

ثبات التجزئة النصفية Split - Half Reliability

إن أبسط أساليب التوافق الداخلي تعرف باسم "التجزئة النصفية" حيث يفصل الاختبار بشكل مصطنع إلى نصفين ويجري ارتباط درجات الأفراد في النصفين. ويطبق الاختبار على مجموعة ثم تقسم البنود بعدئذ إلى نصفين. ويتم الحصول على درجات كل فرد في النصفين ويحسب معامل الارتباط. إن "معامل ثبات التجزئة النصفية" هذا، يشبه معامل التكافؤ لأنه يعكس التذبذب من عينة بنود إلى أخرى. فإذا كان لكل شخص موقع مماثل جداً في الصورتين يكون للاختبار ثبات عالٍ. وإذا كان هناك توافق قليل في المواقع فإن الثبات يكون متدنياً. وتتطلب هذه الطريقة شكلاً واحداً فقط من الاختبار حيث لا يوجد هناك وقت يتبدد، وسوف تعمل التأثيرات البدنية والعقلية على ذات الأشخاص حينما يعالجون النصفين.

أما مشكلة هذه الطريقة فإنها تكمن في فصل الاختبار للحصول على نصفين متماثلين. فإن قام المرء، من خلال تحليل البنود، بتحديد مستوى الصعوبة لكل بند فإن بوسعه وضع كل بند في أحد النصفين على أساس تكافؤ الصعوبة وتشابه المحتوى. أما الإجراء الأكبر شيوعاً، فيتم بربط درجات البنود الفردية الترقيم في الاختبار مع درجات البنود الزوجية الترقيم.

إن معامل الارتباط المحسوب بين النصفين سوف يقلل بشكل منتظم، ثبات الاختبار كله. فالاختبارات الطويلة، أكثر ثباتاً من الاختبارات القصيرة إن كان كل شيء متماثلاً. وعليه فالارتباط بين 50 بنداً فردي الترقيم و 50 بنداً زوجي الترقيم في اختبار ذي 100 بند هو تقدير ثبات اختبار من 50 بنداً وليس من 100 بند. ولأجل تحويل ترابط التجزئة النصفية إلى تقدير ثبات مناسب للاختبار كله، يتم استخدام معادلة سبيرمان – براون (Spearman - Broyn) التنبؤية.

$$r_{xx} = \frac{2r_{1/2\,1/2}}{1 + r_{1/2n1/2}} \qquad (8.5)$$

حيث:

r_{xx} = الثبات التقديري للاختبار برمته.

$r_{1/2\,1/2}$ = معامل ارتباط بيرسون بين النصفين

مثلاً، إذا وجدنا أن معامل الارتباط (0.65) بين نصفي اختبار، فإن الثبات التقديري للاختبار كله، وباستخدام معادلة سبيرمان – براون سيكون:

$$r_{xx} = \frac{(2)(.65)}{1+.65} = .79$$

ويستند إجراء سبيرمان – براون إلى الافتراض القائل بأن النصفين متماثلان. وحيث إن هذا الافتراض قلما يكون صحيحاً بشكل دقيق، فإن الممارسة تفيد أن أسلوب التجزئة النصفية مع تصحيح سبيرمان – براون، يميل إلى المبالغة في تقدير الثبات الذي سوف يتم الحصول عليه باستخدام أساليب إعادة الاختبار أو الصور المتكافئة. فينبغي تذكر ذلك عند تقييم ثبات الاختبارات التنافسية.

يعتبر ثبات التجزئة النصفية أسلوباً مناسباً، إذا توخينا تجنب التذبذب، أحيانا، في تقدير الثبات، وحين يكون الاختبار، نسبياً، طويلاً. وتعد أساليب إعادة الاختبار أو الصور المتكافئة هي الأنسب عندما يكون الاختبار قصيراً.

ولا يعد استخدام التجزئة النصفية مناسباً في اختبارات السرعة لأنه يعطي معاملات تكافؤ عالية في مثل هذه الاختبارات. فاختبار السرعة يشمل عن قصد بنودا سهلة كي تعتمد الدرجات بشكل رئيسي على السرعة التي بوسع الأفراد الاستجابة لها. فالأخطاء صغيرة وأغلب البنود تكون صحيحة حتى نهاية الوقت المحدد. فإن استجاب التلميذ لـ 50 بنداً فيحتمل أن تكون درجة التجزئة النصفية 25–25، وإن أجاب تلميذ آخر على 60 بنداً فيحتمل أن تكون درجة التجزئة النصفية 30–30، وهكذا. وحيث أن درجات الأشخاص في البنود الفردية الترقيم والزوجية الترقيم تتطابق تقريباً، والتباين داخل الأفراد يكون أقل، فإن الارتباط بين الأنصاف سيكون تاماً تقريباً. وهكذا يوصي بأساليب أخرى فيما يخص اختبارات السرعة.

مقاييس التجانس Homogeneity Measures

لا تحتاج مقاييس التوافق الداخلي الأخرى للثبات، فصل الاختبار إلى أنصاف وتصحيح كل نصف بشكل منفصل. فهذه الأساليب تقيّم التوافق ما بين البنود أو "التجانس" لدى البنود. وهي تعكس مصدرين للخطأ: 1) معاينة المحتوى كما في التجزئة النصفية و 2) وانعدام التجانس في المجال السلوكي المعاين. وكلما كان المجال غير متجانس كان التوافق الداخلي للبنود متدنياً، وبعكس ذلك، كلما كان المجال متجانساً كان التوافق الداخلي للبنود عالياً.

أساليب كودر – ريتشاردسون Kuder - Richardson Procedures

لقد قام كودر و ريتشاردسون (1937) بتطوير أساليب استخدمت بشكل واسع لتحديد

التجانس، وربما كان أفضل مؤشر معروف للتجانس هو معادلة كودر - ريتشاردسون 20 التي تقوم على نسبة الاستجابات الصحيحة وغير الصحيحة لكل بند في الاختبار.

$$r_{xx} = \frac{k}{k-1}\left(\frac{S_x^2 - \sum pq}{S_x^2}\right) \qquad \text{k-R 20 (8.6)}$$

حيث:

K = عدد بنود الاختبار.

S_x^2 = تباين درجات الاختبار الكلي (مربع الانحراف المعياري)

P = نسبة الاستجابات الصحيحة لبند واحد.

q = نسبة الاستجابات الخاطئة للبند ذاته.

ويحسب الحاصل pq لكل بند ثم تجمع الحواصل لكل البنود للحصول على Σpq. ويمكن تطبيق معادلة كودر - ريتشاردسون 20 على الاختبارات التي تكون تصحيح بنودها ثنائياً، أي من النوع (صح أو خطأ). وتقدم، نمطياً، العديد من أساليب التصحيح الآلي معامل الارتباط K-R20 مع معامل ارتباط التجزئة النصفية.

وثمة معادلة أخرى (Kuder - Richardson 21) أسهل حسابياً لكنها تتطلب افتراض إن جميع البنود في الاختبار ذات صعوبة متماثلة. وهذا الافتراض غير واقعي في الغالب.

$$r_{xx} = \frac{KS_x^2 - \overline{X}(K - \overline{X})}{S_x^2(K-1)} \qquad \text{K-R 21 (8.7)}$$

حيث:

r_{xx} = ثبات الاختبار كله.

K = عدد بنود الاختبار.

S_x^2 = تباين الدرجات.

$\overline{X}$ = الوسط الحسابي للدرجات.

تعد هذه الطريقة، إلى حد بعيد، الأقل استهلاكاً للوقت من جميع أساليب تقدير الثبات.

وهذا يشمل فقط تطبيقاً واحداً لاختبار ما، ويستخدم فقط معلومات متوفرة يسيرة. ويمكن التوصية به ليستخدمه المعلمون في حجرة الدراسة.

مثلاً، افترض أن مدرساً أجرى اختباراً من 50 بنداً على صف، ووجد الوسط الحسابي (40) والانحراف المعياري (0.6). بتطبيق المعادلة 7.8 يمكن تقدير الثبات كما يلي:

$$r_{xx} = \frac{(50)6^2 - 40(50-40)}{6^2(50-1)} = \frac{1800-400}{1764} = 0.79$$

ونظراً لأن أساليب كودر - ريتشاردسون تؤكد على تكافؤ كل البنود في اختبار معين، فإنها مناسبة بشكل خاص عندما يكون الهدف من الاختبار قياس سمة واحدة. أما إذا كان الاختبار مصمماً لقياس سمات عدة، فإن تقدير ثبات كودر - ريتشاردسون سيكون عادة أدنى من تقديرات الثبات القائمة على إجراء الترابط.

لقد تبين من المحاكمة الاستنتاجية أن ثبات كودر - ريتشاردسون بالنسبة لأي اختبار مكافئ رياضياً لوسط تقديرات ثبات التجزئة النصفية المحسوبة لكل طريقة ممكنة للتجزئة النصفية للاختبار. وتساعد هذه الحقيقة على تفسير العلاقة بين الأسلوبين. فإذا كان الاختبار ذا صعوبة منتظمة ويقيس سمة واحدة، فإن أية طريقة تجزئة نصفية لهذا الاختبار، محتمل لها أن تكون مماثلة لأية تجزئة نصفية أخرى. وعليه، فإن طرق سبيرمان - براون و كودر - ريتشاردسون سوف تعطي تقديرات متماثلة. وإن كان للاختبار بنود متنوعة الصعوبة وتقيس سمات متنوعة، فيتوقع أن يكون تقدير كودر - ريتشاردسون أدنى من تقدير التجزئة النصفية. مثلاً، افترض أن عينة اختبار المهارات السكرتارية، مهارات الطباعة والاختزال والتهجئة والنحو الإنكليزي. وبتطبيق طريقة التجزئة النصفية، فإن مصمم الاختبار سوف يعين أعداداً متساوية البنود من كل اختبار فرعي لكل نصف من الاختبار. فإن كان الاختبار يمضي بشكل جيد في قياس مجموعة المهارات هذه، فإن ثبات التجزئة النصفية سيكون عالياً. أما طريقة كودر - ريتشاردسون التي تقدّر المدى الذي تكون عنده كل البنود متكافئة مع بعضها سوف تعطي تقدير ثبات متدن بشكل كبير.

معامل الارتباط ألفا: هناك مقياس آخر واسع الاستخدام للتجانس هو معامل ألفا، الذي يدعى أيضاً "كرونباخ ألفا" على اسم لي كرونباخ (1951). ولمعامل ألفا استخدامات أوسع من المعادلة K-R20. فحين يتم تسجيل درجات البنود بشكل انقسامي لفئتين / ثنائي فإنه ينتج نتيجة معادلة K-R20 نفسها لكن بالإمكان استخدامه كذلك عندما لا يتم تسجيل الدرجات بشكل انقسامي. أما معادلة ألفا فهي كما يلي:

$$\alpha \text{ or } r_{xx} = \left(\frac{K}{K-1}\right)\left(\frac{S_x^2 - \sum S_i^2}{S_x^2}\right) \qquad (8.8)$$

حيث:

K = عدد بنود الاختبار.

$\sum S_i^2$ = مجموع تباينات البنود.

S_x^2 = تباين درجات الاختبار (جميع بنود K).

إن معادلة ألفا مشابهة لمعادلة K-R20 عدا أن Σpq تستبدل بـــ $\sum S_i^2$، أي مجموع تباينات درجات البنود. ولحساب ذلك يحدد المرء تباين جميع الدرجات لكل بند ثم يجمع تباينات كل البنود للحصول على $\sum S_i^2$. إن ألفا و K-R20 متكافئتان للبنود ثنائية

الاستجابة. وتستخدم معادلة كرونباخ ألفا عندما يكون للمقاييس بنود لا تصحح ببساطة بوصفها صحيحة أو خاطئة كما هو الحال في مقاييس الاتجاهات أو اختبارات المقالة. فقد تأخذ درجة البند واحدة من مدى للقيم كما هو الحال مثلاً، في مقياس ليكرت للاتجاهات حيث قد يحصل الشخص على درجة من 1 إلى 5 استناداً إلى الاختيار الذي يراه. وبصورة مماثلة، ففي اختبارات المقالة قد يخصص عدد مختلف من الدرجات لكل جواب. وسوف تؤمن العديد من برامج الحاسوب الخاصة بالثبات، كتلك الموجودة في برنامج SPSS، المعامل ألفا كمؤشر للثبات.

أما إن كانت بنود الاختبار لا متجانسة – أي أنها تقيس أكثر من سمة – فإن مؤشر الثبات كما يحسب عن طريق معامل ألفا أو K-R20 سوف يتضاءل علاوة على ذلك لا تعد هذه المعادلات مناسبة في الاختبارات الموقوتة طالما أن تباينات البنود ستكون دقيقة فقط إن تم محاولة كل بند من قبل كل شخص.

يعرض الجدول 8.3 خلاصة بالأنماط المختلفة لمعاملات الثبات المرتبة حسب عدد الأشكال / الصور أو عدد تطبيقات الاختبار المطلوبة.

الجدول 8.3: خلاصة لمعاملات الثبات

		عدد صور الاختبار المطلوبة	
		واحد	اثنان
عدد التطبيقات المطلوبة	واحد	التجزئة النصفية K-R20 المعامل ألفا	صور متكافئة (لا إهدار في الوقت)
	اثنان	إعادة اختبار	صور متكافئة (إهدار في الوقت)

تفسير معاملات الثبات

Interpretation of Reliability Coefficients

ينبغي أن يستند تفسير معامل ارتباط الثبات إلى عدد من الاعتبارات. وهناك عوامل معينة تؤثر على معاملات الثبات، وما لم تؤخذ هذه العوامل بنظر الاعتبار، فإن أي تفسير للثبات سيكون غير دقيق.

1. يعتمد ثبات الاختبار، جزئياً على طول الاختبار. فكلما كان الاختبار أطول، كان

ثباته أكبر. فالاختبار يتكون عادة من عدد من بنود عينة تستمد نظرياً من مجتمع بنود الاختبار. ونحن نعرف مما درسناه حول المعاينة بأنه كلما كان حجم العينة كبيراً كان توقعنا بأنها أكثر تمثيلاً للمجتمع الذي تستمد منه. ويصح ذلك أيضاً للاختبارات. وإن كان بالإمكان استخدام بنود المجتمع كلها، فإن درجة الشخص في الاختبار ستكون الدرجة الحقيقية. والمجتمع النظري للبنود يتكون من عدد غير منته من الأسئلة، فهو، بصورة واضحة، لا يشكل احتمالاً عملياً. لذا، فالمرء يبني اختباراً يعتبر عينة لمثل هذا المجتمع النظري. وكلما كان هذا الاختبار أطول (أي كلما كان عدد البنود المشمولة بالاختبار أكبر) كان أكثر تمثيلاً للدرجة الحقيقية للأشخاص الذين يأخذونه. ولأن الثبات هو المدى الذي يمثل فيه الاختبار الدرجات الحقيقية للأشخاص، فإنه كلما كان الاختبار أطول كان الثبات أكبر بافتراض أن جميع بنود الاختبار تنتمي إلى مجتمع البنود.

2. يعتمد الثبات، جزئياً، على عدم تجانس المجموعة. فمعامل الثبات يزداد عندما يزداد انتشار، أو عدم تجانس، الأشخاص الذين يأخذون الاختبار. وبالعكس، كلما كانت المجموعة أكثر تجانساً فيما يخص السمة المقاسة كان معامل الثبات أدنى. واحد تفسيرات الثبات هو أنه المدى الذي يمكن توزيع الأشخاص فيه، بالنسبة إلى الآخرين في مجموعتهم حسب سمات معينة. إن مثل هذا التوزيع، يكون أسهل حين يتعامل المرء مع مقاييس تقع في مدى كبير بدلاً من تلك التي تقع في مدى صغير. فلا يتطلب الأمر أداة حساسة لتحديد وضع الأطفال في توزيع حسب أوزانهم إذا كان المدى العمري للأطفال من 5 إلى 15. وفي الحقيقة، إن هذا الوضع ممكن بدرجة ما من الدقة، حتى بدون استخدام وسيلة قياس. لكن الأمر يحتاج إلى وسيلة قياس حساسة لتنفيذ الوضع ذاته إذا ما كان أعمار كل هؤلاء الذين تتم مقارنتهم ووضعهم في التوزيع، 5 سنوات. وهكذا، فإن عدم تجانس المجموعة التي استخدمت لها أداة القياس هو عامل يؤثر على ثبات الوسيلة. فكلما كان عدم تجانس المجموعة المستخدمة في دراسة الثبات أكبر، كان معامل الثبات أعلى. وينبغي أن لا تغيب عن بالنا هذه الحقيقة عند انتقاء اختبار مقنن. فالناشر قد يذكر معامل ثبات عال يستند إلى عينة ذات مدى واسع للقدرة. وعلى أية حال، عندما يستخدم الاختبار مع مجموعة لها مدى قدرة أضيق كثيراً فإن الثبات سيكون أخفض.

3. يعتمد ثبات الاختبار، جزئياً، على قدرة الأشخاص الذين يأخذون ذلك الاختبار. فقد يكون الاختبار ثابتاً على مستوى قدرة واحد، لكنه لا يكون ثابتاً على مستوى آخر. فأسئلة الاختبار قد تكون صعبة، تتعدى مستوى قدره من يأخذونه أو قد تكون الأسئلة سهلة لغالبية الأشخاص. فمستوى الصعوبة يؤثر على ثبات الاختبار.

فحينما يكون الاختبار صعباً، فإن الأشخاص يخمنون معظم الأسئلة وتكون النتيجة معامل ثبات متدن. وحين يكون سهلاً، فإن استجابات الأشخاص تكون صحيحة في معظم البنود ويميّز القليل من البنود الصعبة بين الأشخاص. ومرة أخرى سوف نتوقع ثباتاً متدن. وليست هناك قاعدة سهلة تتيح لنا تحديد مدى الصعوبة أو السهولة التي ينبغي على الاختبار تحقيقها. إن ذلك يعتمد على نوع الاختبار وهدفه والمجتمع الذي أعد لأجله.

4. يعتمد الثبات، جزئياً، على الأسلوب الخاص المستخدم لتقديره. فالأساليب المختلفة لتقدير ثبات الاختبارات تؤدي إلى معاملات ثبات مختلفة. فأسلوب الصور المتكافئة يعطي تقدير ثبات أدنى من أساليب إعادة الاختبار أو التجزئة النصفية، لأن التذبذب ينشأ في أسلوب الصور المتكافئة من صورة – إلى أخرى ومن وقت – إلى آخر كذلك. أما طريقة التجزئة النصفية، فإنها، من ناحية أخرى، تؤدي إلى معاملات ثبات أعلى من بدائلها، وذلك لأن معظم الاختبارات تنطوي على بعض من درجة السرعة إلى مدى يجعل تقدير معامل الثبات مبالغاً فيه. وهكذا ففي تقييم ثبات اختبار ما، يفضل المرء الاختبار الذي تم تقدير معامل ثباته بأسلوب الصور المتكافئة وليس بالأساليب الأخرى، حينما تكون حالات الثبات المذكورة متشابهة. ويصح التعميم ذاته لدى مقارنة ثبات إعادة الاختبار مع التجزئة النصفية. ويكون معامل الثبات نفسه أكثر إقناعاً إن جاء من إجراء إعادة اختبار وليس من طريقة التجزئة النصفية.

5. يعتمد الثبات، جزئياً، على طبيعة المتغير المقاس. إن بعض المتغيرات ذات الأهمية بالنسبة إلى الباحثين تتمخض عن مقاييس مستقرة أكثر مما تفعله متغيرات أخرى. فمثلاً، لأغلب اختبارات التحصيل الأكاديمي الراسخة ثبات عال جداً، بيد أن لاختبارات متغيرات الشخصية، ثبات متواضع فحسب.

ما هو الحد الأدنى للثبات المقبول لأداة معينة؟ ربما كانت أفضل إجابة على هذا السؤال هو أن الثبات الجيد هو الذي يعتبر جيداً أو أفضل من ثبات المقاييس المنافسة. فاختبار تحصيل التهجئة ذو الثبات (0.80) ليس مرضياً إن كانت لاختبارات منافسة معاملات ثبات تبلغ (0.90) أو أفضل. إن معامل (0.80) لاختبار الإبداع سوف سيعتبر ممتازاً إن كانت للاختبارات الأخرى الخاصة بالمفهوم البنائي ذاته درجة ثبات (0.60) أو أقل.

وتعتمد درجة الثبات المطلوبة في مقياس ما، إلى حد بعيد، على استخدام النتائج المترتبة. فإذا أريد استخدام نتائج القياس لاتخاذ قرار حول مجموعة أو حتى لأغراض بحثية، فإن معامل الثبات المتدني (بمدى 0.60–0.50) قد يكون مقبولاً. غير أنه، إذ أريد

استخدام النتائج كأساس لاتخاذ قرارات حول أشخاص، خصوصاً القرارات المهمة أو التي يتعذر إلغاؤها، عندئذ تكون الأدوات ذات الثبات الأعلى هي المقبولة. لقد أوضح (Nunnally, 1967) أنه في مثل هذه الأوضاع يكون الثبات الذي مقداره (0.90) هو الحد الأدنى الذي ينبغي التسامح معه وإن الثبات البالغ (0.95) ينبغي أن يكون المستوى المفضل.

الخطأ المعياري للقياس Standard Error of Measurement

كما أوضحنا سابقاً في هذا الفصل، فإن الثبات أو الاستقرار أو الموثوقية لاختبار ما، قد يعبر عنها أيضاً في إطار الخطأ المعياري للقياس، الذي يقدم تقديراً لمدى التغير في مجموعة المقاييس المتكررة للشيء ذاته. وبالعودة إلى مثالنا في رمية كرة البيسبول، فقد نتوقع مع الإجراء المتكرر أن نحصل، بالصدفة، على عدد من الدرجات المختلفة لذات الفرد. ونحصل على توزيع تكراري للدرجات. إن لهذا التوزيع التكراري وسطاً حسابياً الذي هو أفضل تقريب للدرجة الحقيقية. كما أن لهذا التوزيع انحرافاً معيارياً يوضح مدى التباين في الدرجات. ولأن هذا الانحراف المعياري هو انحراف معياري لأخطاء القياس فإنه يدعى "الخطأ المعياري للقياس". فإذا أراد المرء إنشاء بناء مضلع تكراري يوضح توزيع الدرجات هذه، فإن شكله سيكون قريباً من المنحنى الاعتيادي / النظامي. وتتوزع أخطاء القياس بشكل اعتيادي، فقد تكون هناك العديد من الأخطاء الصغيرة ولكن هناك القليل من الأخطاء الكبيرة. إن الانحراف المعياري لتوزيع الأخطاء هذا (الخطأ المعياري للقياس، S_M) سيعطينا تقديراً لمدى توقعنا لتكرار حدوث أخطاء ذات حجم معين لدى استخدام الاختبار.

وفي العديد من الأوضاع، لا تتوفر للمرء مقاييس متكررة، ولكن بوسعه أن يحصل على تقدير للخطأ المعياري للقياس باستخدام معامل الثبات:

$$S_M = S_x\sqrt{1-r_{xx}} \qquad (8.9)$$

حيث:

S_M = الخطأ المعياري للقياس.
S_x = الانحراف المعياري لدرجات الاختبار.
r_{xx} = معامل الثبات.

فإذا كان لاختبار الاستعداد معامل ثبات (0.96) وانحراف معياري هو 0.15 فإن

$$S_M = 15\sqrt{1-0.96} = 15\sqrt{0.04} = 3$$

ويعد الانحراف المعياري للقياس مؤشراً للتغير / التشتت المتوقع للدرجات المستحصلة

حول الدرجة الحقيقية. وبالوسع تفسير S_M بكونه الانحراف المعياري لدرجة الخطأ المرتبطة بدرجة مرصودة / ملاحظة ويمكن تفسيره مثل أي انحراف معياري آخر. فإذا أعطينا الدرجة التي أحرزها الطالب، فنستخدم S_M لتحديد مدى قيم الدرجات التي تشمل الدرجة الحقيقية حسب احتمال معين. ويشار إلى مدى الدرجات "فترة / مدى الثقة". وبالوسع افتراض أن أخطاء القياس موزعة بشكل اعتيادي / نظامي حول درجة معينة كما أنها موزعة بشكل متساو عبر مدى الدرجات، فيمكن للمرء أن يكون 68% واثقاً أن الدرجة الحقيقية للشخص تقع ضمن S_M واحد على كل طرف من الدرجة المرصودة. فمثلاً، إذا كان لشخص درجة مرصودة (110) في اختبار للاستعداد حيث يبلغ الخطأ المعياري للقياس (3) فإن المرء يمكن أن يستنتج وبثقة 68%، أن الدرجة الحقيقية للشخص تقع ما بين (107) و (113). أو يمكننا أن نوضح، بمستوى الثقة 95%، أن الدرجة الحقيقية ستقع ضمن $\pm 1.96_{SM}$ للدرجة المرصودة (بين 104 و 116). وبوسعنا أيضاً استخدام الخطأ المعياري للقياس لتحديد مقدار التغير / التشتت المتوقع في إعادة اختبار الشخص. فإذا أمكن إعادة اختبار شخص معين حسب اختبار الاستعداد ذاته، عدداً من المرات فيمكن أن نتوقع أنه في حوالي ثلثي الاختبارات، ستقع الدرجات ضمن مدى (6) درجات وأنه في 95% من درجات إعادة الاختبار ستقع ضمن مدى (12) درجة.

إن الخطأ المعياري للقياس ومعامل الثبات هي طرق متبادلة للتعبير عن ثبات الاختبار، أما مدى دقة تقدير درجة مستحصلة للدرجة الحقيقية فإنه يتضح من خلال حجم الخطأ المعياري للقياس ومعامل الثبات. فالمرء يبحث عن خطأ معياري للقياس يكون متدنياً أو عن معامل ثبات عال كمؤشرات ثبات الاختبار. ويوصى باستخدام الخطأ المعياري للقياس، لتفسير درجات الفرد، ومعامل الثبات لدى مقارنة ثبات / استقرار اختبارات مختلفة.

ثبات الاختبارات مرجعية الإتقان

Reliability of Criterion(*) - Referenced Test

إن تطوير طرق مقنعة لتقدير ثبات الاختبار مرجعي الإتقان لا يزال صعباً. تذكر أن الاختبارات المرجعية الإتقان تستخدم لتحديد حالة الشخص فيما يتعلق بمجموعة معرفة تماماً لأهداف المحتوى / المضمون. إن ثبات هذا النوع من الاختبار يتعلق بالتوافق / الاستقرار الذي تقدر فيه هذه الحالة. فهل يكون للفرد مستوى المهارة نفسها في إجراءين للاختبار؟ إن أساليب الثبات التقليدية، كالارتباط ومعادلة K-R20 لا تعد مناسبة للاختبارات مرجعية الإتقان لأن هذه الأساليب تعتمد على قابلية تغير الأشخاص الذين

(*) المحك أو المعيار (Criterion) يشير في هذا الاختبار إلى "الإتقان" وتحدد شروطه مسبقاً، كما أشرنا فيما سبق - (المراجع)

حسبت لهم. ففي الاختبارات مرجعية الإتقان هناك القليل، أو لا يوجد، تغيرات / تشتت في الدرجات لأن التدريب يستمر حتى يتم إتقان المهارة، أو إتقانها تقريباً. إن التغير المحدود للدرجات أو انتشارها سوف يؤدي إلى تقديرات ثبات متدنية أو شبه صفرية حتى وإن كان الاختبار متوافقاً داخليا ومستقراً بشكل عال.

ولقد تم اقتراح العديد من الأساليب لتقدير ثبات الاختبارات مرجعية الإتقان. فإجراء بسيط نسبياً ينطوي على تطبيق شكلين متكافئين للاختبار أو ذات الاختبار، في مناسبتين وإيجاد توافق / ثبات القرارات التي تم الوصول إليها. ويتحدد توافق النتائج بإيجاد النسبة المئوية للأشخاص الذين يتخذون القرار ذاته (إتقان أو لا إتقان) من أجلهم في كلا التطبيقين. ويشار إلى هذا المؤشر بمعامل الاتفاق (P_o).

فمثلاً، تم الحصول على النتائج المبينة في الجدول 8.4 عندما طُبق شكلان متكافئان من اختبار مرجعي الإتقان على عينة من (100) طالب. وفي مثل هذا الوضع تم تصنيف (70) طالب بشكل متواصل كمتقنين و (14) غير متقنين. ويعتبر معامل الاتفاق هو نسبة جميع الأشخاص الذين صنفوا بشكل متواصل في التطبيقين للاختبار أو:

$$P_o = \frac{a+d}{N} = \frac{70+14}{10} = \frac{84}{100} = 0.84 \qquad (8.10)$$

حيث:

a = العدد المصنف كمتقنين في كلا التطبيقين.
d = العدد المصنف كغير متقنين في كلا التطبيقين.
N = العدد الكلي للأشخاص.
P_o = معامل الاتفاق.

وهكذا فإن 84% من الأشخاص قد صنفوا بشكل متواصل وأن (0.84) هو مؤشر الثبات للاختبار. وإذا كانت التصنيفات للمتقنين أو غير المتقنين متوافقة / مستقرة بالنسبة لجميع الممتحنين في كلا التطبيقين فإن معامل الاتفاق يساوي 1 وهو القيمة القصوى.

الجدول 8.4 القرارات القائمة على الصور 1، 2

		الصورة 1				
		متقن		غير متقن		
الصورة 2	متقن	(a)	70	(b)	10	80
	غير متقن	(c)	6	(d)	14	20
			76		24	100 (N)

إن بعض الاتفاق في التصنيفات متقن – غير متقن ضابط بين تطبيقين يتوقع له أن يكون بمجرد الصدفة، أي حتى، وإن تمت التصنيفات بشكل عشوائي فإن بعض الأشخاص يتوقع أن يقعوا في الفئتين (a) و (d) في المثال. لذا اقترح بعض الكتاب (Swaminathan, Hambleton, & Algina, 1974) استخدام المعامل كابا (K) وهو إحصاءة يأخذ "اتفاق الصدفة" بنظر الاعتبار. ويشير المعامل كابا (K) إلى نسبة التصنيفات المتوافقة التي رصدت، بما يفوق توقع الصدفة.

والأساس المنطقي لمعامل كابا بسيط، فالمرء يحسب أولاً النسبة المئوية للحالات المتوقعة لها أن تكون تصنيفاً متوافقاً حتى إذا لم تكن هناك علاقة أصيلة بين الاختبارات – أي إن كان التصنيف في التطبيقين مستقلاً تماماً. ويشار إلى المؤشر "اتفاق الصدفة المتوقعة" (P_c). ويطرح هذا الاتفاق من الاتفاق المرصود (P_o - P_c) للحصول على الزيادة الفعلية فوق ثبات الصدفة، وتقسم هذه الكمية بعدئذ على ($1-P_c$)، وهي أقصى زيادة ممكنة لثبات القرار فيما يفوق الصدفة للحصول على (K) أو معامل كابا. وهكذا يتضح اتفاق الصدفة المتوقعة بالمعادلة التالية:

$$P_c = \frac{(a+b)(a+c)+(c+d)(b+d)}{N^2} \qquad (8.11)$$

حيث:

P_c = نسبة الاتفاق المتوقع بالصدفة.

ويكون معامل كابا هو:

$$K = \frac{P_o - P_c}{1 - P_c} \qquad (8.12)$$

حيث:

K = نسبة الاتفاق فوق ما هو متوقع بالصدفة.

P_o = معامل الاتفاق المرصود / الملاحظ.

P_c = اتفاق الصدفة المتوقعة.

باستخدام البيانات في المثال أعلاه:

$$P_c = \frac{(80)(76)+(20)(24)}{100^2} = \frac{6080+480}{10,000} = .66$$

$$K = \frac{.84-.66}{1-.66} = \frac{.18}{.34} = .53$$

يمكن أن نرى أن معامل كابا (.53) يعطينا تقديراً للثبات أدنى من معامل الاتفاق (.84) وهذه هي الحال دائماً. باستثناء حالة الاتفاق التام (P_o=1.00) لأن معادلة كابا تبدأ بـ (P_o) ثم تتكيف لاتفاق الصدفة المتوقعة. وتحسب كلتاهما في الغالب لأن المعلومات التي يقدمها المؤشران مختلفة. فمعامل الاتفاق وكابا يتطلبان تطبيقين للاختبار. وقام (Huynh, 1976) و (Subkoviak, 1976) بتطوير أساليب لتقدير هذه المعاملات بإجراء اختبار واحد. وعلى أية حال، فإن أبسط إجراء هو استخدام الجداول المطورة من قبل سابكوفياك (1988) التي تسمح للمرء أن يقرأ القيمة التقريبية لمعامل الاتفاق أو معامل كابا مباشرة بعد إجراء واحد لاختبار مرجعي الإتقان.

ثبات البيانات الملاحظة / المرصودة

Reliability of Observational Data

يعتبر الثبات مهما في أدوات القياس التي تتطلب تقديرات أو ملاحظات للأشخاص من أشخاص آخرين. إن على الباحث في هذه الحالات أن يحدد ثبات التقديرات – ما إذا أعطى حكام/ ملاحظون مختلفون درجات أو تقديرات متماثلة للسلوك ذاته. وكطريقة بسيطة لتحديد ثبات التقديرات تكون بقيام مراقبين أو أكثر بتقدير السلوك ذاته بصورة مستقلة ثم دراسة تقديرات المراقبين / الملاحظين بعدئذ. إن الارتباط الناجم يدعى "الثبات الداخلي للتقديرات أو المراقبين"، فإذا كانت أنواع السلوك المراد رصده أو ملاحظته محددة بشكل جيد وكان الملاحظون مدربين بشكل جيد، فينبغي أن يكون ثبات الملاحظات موجباً وعالياً جداً.

خذ حالة شخصين قاما بتقدير عدة طلبة في تقييم أداء حيث كانت التقديرات تتراوح من 1 (ضعيف جداً) إلى 10 (ممتاز). هنا يمكن تقدير الثبات من خلال إجراءات ترابطية بنفس الطريقة التي تستخدم فيها هذه الإجراءات في الاختبار إعادة اختبار أو ثبات الأشكال / الصور المتكافئة. إن المراقب الثاني يؤدي مهمة إعادة اختبار أو الشكل المكافئ في اختبار القلم والورقة.

وحين تكون الدرجات فقط (1) أو صفر (السلوك الذي حدث مقابل السلوك الذي لم يحدث) فإن معامل الاتفاق (المعادلة 8.10) أو كابا (المعادلة 8.12) يمكن استخدامهما لتقدير ثبات درجات المراقبين.

وتعتبر هذه الإجراءات مفيدة لدى تدريب المراقبين. فالمتدربون يراقبون شريط فيديو ويضعون درجات سبق لمراقب خبير وضع الدرجات الخاصة بالشريط، ويوضح معامل الاتفاق أو كابا التوافق / التطابق بين المتدرب والمراقب الخبير، فالمدرب يقوم بتدقيق الشريط مع المتدرب وذلك لتحديد متى ولماذا أساء المتدرب تصنيف الملاحظات.

ويمكن إيجاد مناقشة واسعة لأساليب الثبات في حالات الرصد السلوكي في كتاب (Suen & Ary, 1989).

مقارنة الصدق والثبات

Validity And Reliability Compared

يعد الصدق، خاصة أكثر أهمية وشمولا من الثبات. كما أنه أكثر صعوبة في التحديد. وتذكر أغلب الدراسات البحثية المنشورة بيانات الثبات أكثر بكثير من بيانات الصدق. وهذا يعزى إلى كون أن الصدق لا يمكن الحصول عليه بشكل مباشر كالثبات. وتقدير أداة قياس يتطلب حكما ذاتياً / غير موضوعي بدرجة كبيرة. وعلى المرء أن يجيب على أسئلة حول ملاءمة محتوى الاختبار، وكفاية المعايير، وتعريفات السمات الإنسانية، ومواصفات المجال السلوكي، والنظرية المتعلقة بمحتوى الاختبار، وما إلى ذلك. فجميع هذه القضايا تشمل الحكم وجمع البيانات من العديد من المصادر.

من الناحية الأخرى، يمكن التحقيق في الثبات بشكل مباشر من بيانات الاختبار، ولا حاجة لبيانات خارجية. إن القضايا الأساسية للثبات تلائم، بسهولة، التحليل الرياضي والاستنتاجات المنطقية حول مقدار الخطأ الذي يمكن التعبير عنه بصورة رياضية.

ولأجل أن يكون المقياس صادقاً، فلا بد له أن يتمتع أيضاً بالثبات. فثبات أداة للقياس يحدد الحد الأعلى لصدقها. أما درجات اختبار ذي ثبات صفري فهي، كلياً، عشوائية، ولا يمكن دراسة ترابطها مع أي معيار.

والترابط المحتمل للأداة مع أي معيار ممكن، يزداد مع زيادة ثبات الأداة. ومع ذلك، تذكر أنه يمكن للمقياس أن يتمتع بالثبات دون أن يكون صادقاً. وبوسعه أن يقيس بشكل متواصل الشي الخطأ. وهنا يؤكد (Feldt & Brennan, 1989, p. 143) على أولوية الصدق في تقييم كفاية المقياس التربوي بقولهما: "إن أي كيان لبيانات الثبات، وبغض النظر عن كفاءة الطرق المستخدمة لتحليلها، لا تستحق الكثير، إن كان المقياس المطبق غير مناسب أو فائض".

الخلاصة SUMMARY

إن الكثير من أدوات القياس المتيسرة لدى الباحث تتطلب استخدام معايير لتقييم هذه الأدوات. والمعياران المهمان لأدوات القياس هما الصدق والثبات.

يشير الصدق إلى ملاءمة ووضوح وفائدة الاستدلالات القائمة من درجات الاختبار. ففي عملية تحقيق صدق اختبار، يستخلص المرء الدليل لدعم الاستنتاجات.

وثمة إجراءات متنوعة لاستخلاص دليل حول صدق الاختبار. ويشير الدليل المتعلق بالمحتوى إلى مدى معاينة محتوى الاختبار بمجال موضوع الدراسة الذي تستخلص النتائج من أجله. كما يشير الدليل المتعلق بالمعيار إلى مدى ترابط درجات الاختبار مع معايير مناسبة. أما الدليل المتعلق بالمفهوم البنائي فيوضح المفاهيم البنائية النفسية أو التربوية التي يقيسها الاختبار.

ومن المهم معرفة نوع الاستخدام لدرجات الاختبار. فقد يكون الاختبار صادقاً في استخدام ما ولكنه غير صادق في استخدامات أخرى. فالاختبار المصمم لقياس الاستعداد الدراسي قد يكون مقياساً صادقاً للموهبة الأكاديمية، ولكنه غير صادق كمقياس للاستعداد الفني.

ويجب أن نسأل كذلك، ما مدى توافق / استقرار ما يقيسه الاختبار، أيا كان ما يقييس؟ هذه هي قضية الثبات. فلا يمكن أن يكون اختبار ما صادقاً ما لم تكن قياساته متوافقة- أي ما لم يتمتع بالثبات. فالثبات يشير إلى المدى الذي يكون فيه الاختبار متوافقاً في قياس ما يقيسه. وبصورة محددة، يشير الثبات إلى المدى الذي تكون فيه درجات الشخص متماثلة تقريباً في مقاييس متكررة كما هو موضح عن طريق معامل ثبات عال. ويمكن حساب معاملات الثبات بطرق مختلفة، استناداً إلى مصدر الخطأ الذي أخذ بنظر الاعتبار. ويبين معامل الثبات المدى الذي تؤثر فيه الأخطاء العشوائية للقياس على درجات الاختبار. والخطأ المعياري للقياس يمكّننا من استخدام المنحنى الاعتيادي / النظامي لتقدير الحدود التي يتوقع أن تقع ضمنها الدرجة الحقيقية للشخص.

ونحتاج إلى أنواع مختلفة من معاملات الثبات للاختبارات مرجعية المتوسط ومرجعية الإتقان، وتأكيد ثبات الأخير أكثر أشكالاً، لكن تطبيقين يستخدمان للاختبارات مرجعية الإتقان قد عرضا في هذا الفصل.

مفاهيم أساسية Key Concepts

معامل اتفاق	agreement coefficient
معامل ألفا	coefficient alpha
(دليل) الصدق المتعلق بالمفهوم البنائي	construct-related validity (evidence of)
(دليل)الصدق المتعلق بالمحتوى	content-related validity (evidence of)
(دليل) الصدق المتعلق بالمعيار	criterion-related validity (evidence of)

الصورة المكافئة	equivalent-form
توافق داخلي	internal-consistency
معادلات كودر – ريتشاردسون	Kuder-Richardson formulas
الثبات	reliability
معادلة سبيرمان – براون	Spearman-Brown formula
ثبات التجزئة النصفية	split-half reliability
الخطأ المعياري للقياس	standard error of measurement
ثبات إعادة الاختبار	test-retest reliability
الصدق	validity

تماريـــن EXERCISES

1. قارن "الصدق" و "الثبات" فيما يخص الآتي:
   أ. معنى كل مفهوم
   ب. الأهمية النسبية لكل مفهوم.
   جـ. المدى الذي يعتمد فيه الواحد على الآخر.
2. أوضح العبارة: قد تكون أداة قياس ثابتة دون أن تكون صادقة، غير أنه لا يمكن لها أن تكون صادقة دون أن تكون ثابتة.
3. كيف تقترح إضفاء الصدق على اختبار جديد للاستعداد الدراسي الذي تم تطويره لطلبة الصفوف المنتهية في الثانوية؟
4. طلب منك إضفاء الصدق على أداة مصممة لقياس المفهوم الذاتي الأكاديمي لطالب (أي الطريقة التي يرى فيها نفسه كطالب). ماذا تفعل لتأكيد صدق هذه الأداة؟
5. أي من الأنماط الثلاثة للدليل على الصدق موضحة في كل من المواقف التالية:
   أ. مقارنة درجات اختبار المهارة اللغوية في الثانوية للمتسربين والمستمرين في الكلية بغية تحديد ما إذا كانت بيانات الاختبار مرتبطة مع حالات طلبة الكلية.
   ب. وجد إن لاختبار جديد للاستعداد الدراسي ارتباط مقداره (93.0) مع SAT الذي يجرى استخدامه للتنبؤ بالنجاح في الكلية.
   جـ. لقد تم تطوير اختبار جديد للذكاء. ويذهب المؤلف إلى أن العمليات العقلية التي يقتضيها الاختبار تتطابق مع نظرية (Z) للذكاء. علاوة على ذلك، يبين المؤلف أن الدرجة المتوسطة في الاختبار بين الأطفال، يزداد مع كل سنة عمرية.

د. يتفحص مدرس بدقة اختبار تحصيل مقنّن ليرى ما إذا كان يغطي المعرفة والمهارات المؤكد عليها في الصف.

هـ. متوسط الفرق بين مراتب أفراد عصابة كوكلاس كلان وأفراد من جماعة الأمريكيين للعمل الديمقراطي حسب مقياس التحررية قد وجد عالياً جداً.

و. الحكم على اختبار الرياضيات من مجموعة مدرسين فيما إذا كان عينة كافية وممثلة لمجتمع بنود الاختبار.

6. حدد نوع الإجراء لتقدير الثبات في كل مما يلي:

أ. أعطي الاختبار نفسه مرتين لمجموعة معينة. وكان الارتباط بين الدرجات في الإجرائين للاختبار هو (.90).

ب. كان الترابط بين درجات المجموعة في البنود الفردية لاختبار، والبنود الزوجية للاختبار نفسه: ($r_{xx} = 0.95$).

جـ. أشكال متوازية متكافئة للاختبار قد أجريت بعد شهر واحد وتم ربط نتائج الاجرائين: ($r_{xx} =.86$).

د. جرى استخدام التباين والوسط الحسابي وعدد البنود لتقدير الثبات.

7. كيف تفسر الفروقات في معاملات الثبات في التمرين 6، على فرض أن المجموعات التي أختبرت كانت نفسها.

8. كيف تضفي الصدق على اختبار الاستعداد للقراءة؟

9. ماذا يمكن أن يفعل المرء لزيادة الثبات عند بناء اختبار ما؟

10. أوضح نوع الدليل الذي يعتبر الأكثر أهمية لإضفاء الصدق على الأنواع التالية من الاختبارات:

أ. اختبار التهجئة الصفي.

ب. أداة لقياس دافعية التحصيل.

جـ. قياس مصمم لتحديد المتسربين المحتملين.

11. أوضح كيف يمكن لاختبار الرياضيات أن يكون ذا صدق عالٍ في أحد صفوف الرياضيات وصدق متدن في صف آخر.

12. انتقد العبارة التالية: ثبات اختبار الذكاء هو (0.90)، وعليه يمكن للمرء أن يفترض أن الاختبار يقيس الذكاء فعلياً.

13. حدد الخطأ المعياري للقياس في اختبار ذي انحراف معياري (16) ومعامل ثبات (r_{xx}=.84). كيف تفسر هذا الخطأ المعياري للقياس؟

14. انتخب اختبار تحصيل مقنّن قد تستخدمه في دراسة بحثية، وتحصل على بيانات الصدق الضرورية لهذا الاختبار. (قد تستخدم Buros وكراسة التعليمات التي تصاحب الاختبار الذي تختاره).

15. دقق كراسة تعليمات الاختبار لاختبار تحصيل يستخدم في مدرستك. ما نوع بيانات الثبات المذكورة هناك؟

16. تم الحصول على البيانات التالية عندما أعطي شكلان / صورتان لاختبار مرجعي الإتقان في الرياضيات إلى مجموعة من أطفال مدرسة ابتدائية. وكان هناك 50 بنداً في كل شكل. ولأجل النجاح يتعين على الطالب أن تكون إجابته صحيحة في 80% من بنود كل شكل. عبر عن ثبات هذا الاختبار بصيغة معامل الاتفاق (P) ومعامل كابا (K).

ممتحن	شكل 1	شكل 2
1	45	47
2	43	48
3	45	31
4	39	39
5	39	48
6	34	37
7	46	46
8	48	49
9	43	38
10	36	46
11	45	48
12	38	39
13	44	45
14	31	34
15	42	48

17. انتقد إجراءات تحقيق الصدق التالية:

أ. طور مدرس اللغة الإنكليزية في ثانوية اختبار كتابة لتحديد طلبة الثانوية الموهوبين وأجرى الاختبار على طلبته في الصفوف النهائية للإنجليزية. وعلى أساس الدرجات العالية، سمح للطلبة التسجيل في صف للإنجليزية في جامعة محلية. وفي نهاية الفصل الدراسي، قام المدرس بدراسة الترابط بين درجات الاختبار الأصلي مع تقديرات الطلبة في الإنجليزية بالكلية. وقد دهش المدرس عندما وجد ترابطاً ضئيلاً جداً. فما هي المشكلة؟

ب. طور مستشار المدرسة مقياساً لقياس الحاجة للتحصيل الأكاديمي لدى أطفال مدرسة ابتدائية. وقد طبق المقياس على صفين من أطفال الابتدائية وأعطيت النتائج إلى مدرسي هؤلاء الأطفال. وطلب من المدرسين أن يراقبوا هؤلاء الأطفال بدقة لفصل دراسي واحد، وأن يعطوا تقديرات للأطفال حسب حاجتهم للتحصيل. وتم ربط تقديرات المدرسين بعدئذ مع الدرجات التي حصل عليها الأطفال في المقياس. وكان الارتباط عالياً، فاستنتج المستشار أن للمقياس صدقاً عال لقياس الحاجة للتحصيل. هل تتفق مع استنتاج المستشار؟

18. افترض إنك أردت التحقيق في "إنهاك" المعلمين. اقترح بعض المؤشرات لهذا المفهوم البنائي التي قد تستخدمها لقياس هذا المفهوم البنائي.

19. أجري اختبار "صح – خطأ" على عشرة طلبة. أن "1" يشير للإجابة الصحيحة بينما "صفر" يشير للإجابة غير الصحيحة. وكانت النتائج كما يلي:

	البنود										
الطلبة	1	2	3	4	5	6	7	8	9	10	المجموع
أ	1	1	1	1	1	1	1	1	0	0	8
ب	1	1	1	1	0	1	0	1	0	0	6
جـ	1	1	1	1	1	0	0	0	0	0	5
د	1	1	1	0	0	0	0	1	1	0	5
هـ	1	1	1	1	1	1	1	1	0	0	8
و	1	1	1	1	1	1	1	1	1	1	10
ز	1	1	1	1	1	1	1	1	1	0	9
ح	1	1	1	0	0	0	0	0	0	0	3
ط	1	1	1	1	0	0	0	0	0	0	4
ي	1	1	1	1	0	1	1	1	1	1	9

أحسب معامل ثبات التجزئة النصفية لهذا الاختبار وذلك بتجزئة الاختبار إلى مجموعتي بنود فردية، وبنود زوجية.

20. ما نوع تقدير الثبات الأكثر ملاءمة لأدوات القياس التالية؟

أ. اختبار تحصيل لاختيار من متعدد سوف يستخدم كمتغير تابع في دراسة تجريبية.

ب. سيدرس باحث التغيرات في الاتجاه وسوف يطبق شكلاً واحداً من مقياس للاتجاهات كقياس قبلي وقياس بعدي.

جـ. لدى باحث صورتان لاختبار تحصيل، حيث طبق واحدة منهما في بداية الدراسة والأخرى عند نهاية الدراسة، وهو يريد تحديد ثبات الاختبار.

21. قسم اختبار من 100 بند إلى نصفين ووجد إن معامل ارتباط التجزئة النصفية هو(‎.60‎). أحسب معامل الثبات للاختبار كله.

22. باستخدام مقياس من (10) درجات أعطى حكمان التقديرات التالية لمقالات كتبها مجموعة طلبة. أحسب مؤشراً يوضح ثبات إجراء هذا التقدير.

	حكم1	حكم2
ماري	10	9
جون	8	7
بيتر	7	10
سوزان	9	8
جين	6	5
مارك	4	3

الأجوبة ANSWERS

1. يشير الصدق إلى المدى الذي تقيس فيه أداة، شيئاً، جرى تصميمها لقياسه. أما الثبات فهو المدى الذي تكون فيه أداة متوافقة / ثابتة في قياسها. ويعتبر الصدق مظهراً أكثر أهمية من الثبات لأن الافتقار إلى الصدق يتضمن افتقارا للمعنى. وعلى أية حال، لا يمكن لأداة أن تكون صادقة دون أن تكون، أولاً، ثابتة.

2. قد يقدم مقياس معين درجات متوافقة (ثبات)، لكنه قد لا يحمل أية صلة بمقاييس أخرى مقبولة للمفهوم البنائي، أو قد لا يكون قادراً على التنبؤ بالسلوك المرتبط بالمفهوم البنائي (الصدق). إن درجات اختبار مع ثبات مقداره (صفر) هي درجات عشوائية تماماً، وعليه لا يمكن أن تترابط مع أي معيار. إن مدى الثبات يقدر حداً أعلى لصدق ممكن.

3. لتحديد صدق المفهوم البنائي، يجب علينا أولاً أن نحدد معنى الاستعداد. فإذا أراد أحدهم في قياس القدرة الأكاديمية العامة، فمن الممكن تحديد صدق المحتوى بتفحص بنود الاختبار لتوضيح تمثيلها. فهل قام بتقييم المهارات الأكاديمية الأساسية للقراءة

والتهجئة، والرياضيات، وما إلى ذلك؟ ويمكن تقدير الصدق المرتبط بالمعيار عن طريق معاملات الارتباط بين درجات الاختبار ومعدل النقاط "GPA" للسنة الأخيرة و "GPA" للسنة الأولى، وغيرها من المعايير. ويمكن أيضاً إجراء ارتباط مع درجات اختبار للاستعداد جرى تحقيق صدقه.

4. تحتاج بنود المقياس أو الاستبيان إلى تغطية جوانب سلوك الطلبة التي تكون منطقياً، جزءاً من المفهوم الذاتي الأكاديمي (مثلاً أنا أنوي الذهاب إلى الكلية). وقد تكون مقاييس المعيار على شكل مقابلات شخصية مع الطلبة أو تقدير مستقل من قبل المدرسين. وبافتراض أن المفهوم الذاتي الأكاديمي يرتبط بالتحصيل، فإن درجات المفهوم الذاتي قد تترابط مع GPA و/أو درجات اختبار التحصيل.

5. أ. مرتبط بالمعيار.
ب. مرتبط بالمعيار.
جـ. مرتبط بالمفهوم البنائي.
د. مرتبط بالمحتوى.
هـ. مرتبط بالمفهوم البنائي.
و. مرتبط بالمحتوى.

6. أ. ثبات إعادة الاختبار
ب. ثبات التجزئة النصفية.
جـ. ثبات الأشكال المتكافئة.
د. تكافؤ منطقي (معادلة كودر – ريتشاردسون 21).

7. يميل ثبات التجزئة النصفية إلى أن يكون أعلى من ثبات إعادة الاختبار، لأن تغير الشخص الذي يعزى إلى النضج، والزيادة في مهارة معالجة الاختبار والعوامل العشوائية الأخرى يكون أقل. ويعد ثبات الأشكال المتكافئة أدنى من ثبات الاختبار نفسه لأنه (أ) يستحيل بناء أشكال متكافئة تماماً و (ب) هناك مصدر مضاف للتغير عندما تستخدم أشكال غير متطابقة. أما صدق التكافؤ المنطقي فسوف يقل إن لم يكن الاختبار متجانساً.

8. يحدد المرء أولاً المهارات الخاصة (مثلاً، تمييز الحرف، الاتجاه من اليمين إلى اليسار، بالعربية) التي تكوّن الاستعداد للقراءة، ثم يحدد ما إذا كان الاختبار يدمج هذه المهارات بنسب مناسبة. فعندما يبدأ الأشخاص، الذين أخذوا الاختبار، ببرامج القراءة، فيمكن للمرء أن يحدد الكيفية التي ترتبط فيها درجات الاختبار

والاختبارات الفرعية، مع درجات اختبار القراءة وتقديرات المدرسين والمعايير الأخرى.

9. إعادة كتابة البنود الغامضة وإيضاح التعليمات، سوف يزيد من الثبات. وبإطالة الاختبار عن طريق إضافة بنود مستمدة من نفس المجتمع الإحصائي يزداد الثبات تماماً، كما هو الحال لدى اختبار مجموعة يكون عدم تجانسها أكبر.

10. أ. المحتوى.

ب. مفهوم بنائي.

جـ. مرتبط بالمعيار.

11. إن اختبار الرياضيات الذي يغطي العمليات الحسابية فحسب، سيكون صدقه ضئيلاً في صف يشدد على المفاهيم والتفكير المنطقي. وإذا كان المحتوى والتأكيد في صف مختلف يتفق مع محتوى وتأكيد الاختبار فسوف يكون للاختبار صدق عال في ذلك الصف.

12. يمكن أن يكون يتمتع اختبار بالثبات دون قياس ما ينوي قياسه. ولتحديد الصدق يحتاج المرء إلى أن يتفحص المحتوى، والمفاهيم البنائية، والعلاقات مع مقاييس أخرى للمفهوم البنائي نفسه والعلاقات مع مقاييس للسلوك التي يفترض أنه سيجرى دراسة ترابطها مع المفهوم البنائي.

13. عن طريق المعادلة 8.9

$$S_M = S_x\sqrt{1-rxx}$$
$$= 16\sqrt{1-.84}$$
$$= (16)(.4)$$
$$= 6.4$$

ويفسر المرء خطأ القياس المعياري على أنه انحراف معياري. وهكذا، يمكن أن نقول إن هناك فرصتين من ثلاث بأن الدرجة الحقيقية للشخص سوف تقع في المدى 6.4± درجة من الدرجة المرصودة / الملاحظة.

14. ستختلف الأجوبة.

15. ستختلف الأجوبة.

16. الدرجة 40 تمثل الإتقان ($50 \times 0.80 = 40$).

الصورة 2		الصورة 1: متقن		غير متقن		
متقن	(a)	7	(b)	2		9
غير متقن	(c)	2	(d)	4		6
		9		6		15

$$P_o = \frac{7+4}{15} = \frac{11}{15} = .73$$ (73% من الطلبة قد صنفوا بشكل متواصل / ثابت)

$$P_c = \frac{(9)(9)+(6)(6)}{15^2} = \frac{81+36}{225} = \frac{117}{225} = .52$$

$$K = \frac{.73-.52}{1-.52} = \frac{.21}{.48} = .44$$

17. أ. اختيار درجات عالية فحسب، يقيد التغير. وقد قلص التغير المقيد من معامل الارتباط.

ب. كان هناك إفساد معياري. فإتاحة الفرصة للمدرسين لرؤية نتائج القياس الأصلي للحاجة إلى التحصيل قد أفسد تقديرات الأطفال حول حاجتهم للتحصيل.

18. هناك عدد من المؤشرات المحتملة لإنهاك المدرس. وبوسع المرء أن ينظر إلى الغياب عن المدرسة، وتقييمات المشرفين المتدنية، وحالات العداء نحو التلاميذ أو المشرفين، وحالات تحدي سياسات المدرسة. وقد يطور أحدهم مقياساً لقياس الاتجاهات نحو عملهم، ومن استجابات المدرسين لأسئلة مناسبة، قد يستنتج المرء وجود الإنهاك.

19.

الطالب	فردي	زوجي	المجموع
أ	4	4	8
ب	2	4	6
ج	3	2	5
د	3	2	5
هـ	4	4	8
و	5	5	10
ز	5	4	9
ح	2	1	3
ص	2	2	4
ي	4	5	9

الترابط بين البنود الفردية والزوجية هو (r = .72) (أحسب معامل الارتباط (r) باستخدام معادلة الدرجة الخام مع (N) في المقام)، وباستخدام معادلة سبيرمان - براون يكون تقدير ثبات الاختبار الكامل هو (0.84).

20. أ. قد يهتم المرء بالثبات الداخلي لهذا الشكل من الاختبار. وستكون التجزئة النصفية، أو ألفا أو معامل الثبات حسب كودر - ريتشاردسون مناسباً.

ب. باستخدام شكل واحد كمقياس قبلي ومقياس بعدي سيحسب المرء معامل الثبات.

جـ. مع شكلين واجرائين سيحسب المرء معامل الثبات / الاستقرار والتكافؤ.

21. $r_{xx} = \frac{2(.60)}{1+.60} = .75$

22. بيرسون r = .78

سبيرمان rho = .66 (إنها غير متطابقة بسبب حالتي الرتب المتساوية).

REFERENCES المصادر

American Psychological Association. (1985). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Buros, O.K. (1949). *Mental measurements yearbook* (Vol. 4). Lincoln: University of Nebraska Press.

Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait - multimethod matrix. *Psychological Bulletin,* 56, 81-105.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,* 16,297-334.

Doll, E.A. (1935, 1949, 1965). Vineland social maturity scale. Circle Pines, MN: American Guidance Service.

Feldt, L.S., and Brennan, R.L. (1989). Reliability. In R. Linn (Ed.), *Educational measurement* (pp. 105-146). New York: American Council on Education and Macmillan.

Huynh, H. (1976). On the reliability of decisions in domain-referenced testing. *Journal of Educational Measurement,* 13, 265-266.

Ruder, G.F., and Richardson, M. W. (1937). The theory of estimation of test reliability. *Psychometrika,2,* 151-160.

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement (pp. 13-103). New York: American Council on Education and Macmillan.*

Nunnally, j.C. (1967). *Psychometric theory.* New York: McGraw-Hill.

Subkoviak, M.j. (1976). Estimating reliability from a single administration of a mastery test. *Journal of Educational Measurement,* 13, 265-276.

Subkoviak, M.j. (1988). A practitioner's guide to computation and interpretation of reliability indices for mastery tests. *Journal of Educational Measurement,* 25; 47-55.

Suen, H.R., and Ary, D. (1989). *Analyzing quantitative behavioral observation data.* Hillsdale, Nj: Lawrence Erlbaum.

Swaminathari, H., Hambleton, R.R" and Algina, *j.j.* (1974). Reliability of criterion-referenced tests: A decision-theoretic *formulation. Journal of Educational Measurement,* 11, 263-267.



طـرق البحـث

RESEARCH METHODS

* البحث التجريبي.
* البحث العليّ - المقارن.
* البحث الترابطي / الارتباطي.
* البحث المسحي.
* البحث النوعي والتاريخي.



البحث التجريبي

Experimental Research

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون بوسع الطالب أن:

1- يصف خواص البحث التجريبي.

2- يميز بين المتغير المستقل ومستويات المتغير المستقل.

3- يفسر مفهوم الضبط / التحكم (control) وأهميته في البحث التجريبي.

4- يميز بين المجموعة الضابطة الحقيقية ومجموعة المقارنة في البحث التجريبي.

5- يعرف الصدق الداخلي والخارجي.

6- يحدد التهديدات للصدق الداخلي والاستراتيجيات لتجنبها أو تقليصها إلى الحد الأدنى.

7- يحدد مشكلات الصدق الداخلي في التقارير البحثية.

8- يحدد التهديدات للصدق الخارجي والاستراتيجيات لتجنبها.

9- يحدد مشكلات الصدق الخارجي في المقترحات / الخطط والتقارير البحثية.

10- يصف الفرق بين الاختيار العشوائي والتخصص العشوائي ويبين كيفية ارتباطهما بالصدق الداخلي والصدق الخارجي.

11- يصف التصميمات التجريبية الحقيقية الأكثر شيوعا ويبين فوائد كل منها.

12- يعرف التصميم العاملي ويبين فائدته.

13- يصف البحث التجريبي للشخص الواحد واستخداماته.

تعتبر التجربة استقصاءا علمياً، يغير / يفعّل الباحث فيه واحداً أو أكثر من المتغيرات المستقلة، ويعمل على ضبط المتغيرات ذات الصلة، ويرصد أثر التغيير / التفعيل على المتغير التابع. ويعرض المجرب بشكل مدروس ومنتظم التغيرات، ثم يرصد نتائجها. ولأن التجربة تنطوي على الضبط والرصد الدقيق والقياس فإن طريقة البحث هي التي تقدم الدليل الأكثر إقناعاً لتأثير أحد المتغيرات على آخر.

لقد تعلم العلماء الأوائل قيمة الرصد / الملاحظة في دراسة بيئتنا، لكنهم أدركوا بسرعة إن تعقيد الطبيعة لا يمكن فهمها دائماً من خلال الرصد البسيط لأحداثها العديدة. وقد وجدوا إن الأحداث التي تحدث في حالتها "الطبيعية" قد عقدتها في الغالب عوامل لا صلة لها بحيث إن عملية العامل الذي رغبوا في دراسته قد أصبحت غامضة. وتم حل الصعوبة بضبط الظروف التي حدثت في ضوئها، وذلك لإقصاء العوامل التي لا صلة لها. وعندئذ يمكنهم بشكل مدروس استخدام المتغيرات المعنية المستقلة وقياس التغيرات في المتغيرات التابعة الناتجة عن التغيرات في المتغيرات المستقلة وبعبارة أخرى، لقد بدءوا في إجراء التجارب.

ونظراً لأن تطبيق الطرق التجريبية كان مثمراً في دراسة العالم المادي / الطبيعي فقد طبقت هذه الطرق على ميادين أخرى. وشهد القرن التاسع عشر إدخال هذه الطرق إلى العلوم البيولوجية وتم إحراز تقدم كبير في علم الحيوان والفسيولوجيا والطب. وزهاء نهاية القرن التاسع عشر بدأ الباحثون بتطبيق الطرق ذاتها على المسائل النفسية، وهكذا بدأوا بعلم النفس التجريبي. ففي التسعينات من القرن التاسع عشر استخدمت أول طريقة تجريبية لدراسة مسألة تربوية. وكان بحث (rice, 1897) في تحصيل التهجئة في المدارس يعد أول محاولة بارزة في التجريب في الميدان التربوي. وقام (Thorndike, 1924) وآخرون بتوسيع الطريقة التجريبية على التربية.

فللتجربة في أبسط أشكالها، ثلاث خواص: 1) تغيير / تفعيل المتغير المستقل. 2) إبقاء جميع المتغيرات الأخرى، عدا المتغير المستقل، ثابتة. 3) يرصد أثر تفعيل المتغير المستقل على المتغير التابع. وهكذا، ففي تجربة معينة يكون المتغيران الأكثر أهمية هما المتغير المستقل والمتغير التابع. أما "المتغير المستقل" فيفعّل (يغيّر) من قبل المجرب. ويدعى المتغير الذي ترصد آثار التغيرات عليه بالمتغير التابع وهو يرصد دون أن يفعّل من قبل المجرب. ودعي المتغير التابع بهذا الاسم لأنه يفترض بأن قيمته تعتمد على قيمة المتغير المستقل وتتغير معها. فمثلاً، لأجل دراسة أثر طرق التدريس المختلفة على التحصيل في القراءة يقوم الباحث بتفعيل / تغيير الطريقة (متغير مستقل) باستخدام طرق تدريس مختلفة بغية التأكد من آثارها على التحصيل (متغير تابع).

خواص البحث التجريبي

CHARACTERISTICS OF EXPERIMENTAL RESEARCH

هناك ثلاثة مكونات أساسية يعمل معها العالم بنشاط لدى إجراء أية تجربة وهي: الضبط والمعالجة / التفعيل والرصد / الملاحظة.

الضبط Control

يعد الضبط، جوهر الطريقة التجريبية. فبدون الضبط، يستحيل تقييم آثار المتغير المستقل بشكل لا لبس فيه أو الخروج باستنتاجات حول العلّية. دعنا نتفحص مفهوم الضبط في التجربة، بشكل موجز.

تستند، أساساً، طريقة العلوم التجريبية على افتراضين فيما يخص المتغيرات (Mill, 1846): إذا تكافأ / تماثل موقفان في كل جانب عدا المتغير الذي يضاف أو يحذف من أحد الموقفين فإن أي اختلاف يظهر بين الموقفين يمكن أن يعزى إلى ذلك المتغير. ويشار إلى هذه العبارة باسم "قانون المتغير المستقل الواحد". 2) إذا لم يتكافأ موقفان، غير أنه بالوسع تأكيد أنه يتعذر لأي من المتغيرات، عدا المتغير المستقل، أن يكون متميزاً في إحداث الظاهرة قيد الدرس، أو إذا أمكن إقامة تكافؤ بين متغيرات متميزة، عدا المتغير المستقل، فإن أي اختلاف يقع بين الموقفين بعد إدخال متغير جديد (متغير مستقل) إلى أي من النظامين يمكن أن يعزى إلى المتغير الجديد، وتدعى هذه العبارة "قانون المتغير الواحد المتميز".

إن هدف الضبط في أية تجربة هو ترتيب موقف يمكن أن يدرس فيه أثر المتغيرات. ويحتمل الإيفاء بالظروف، التي تشكل الأساس لقانون المتغير الواحد، في العلوم الطبيعية أكثر مما في التربية. مثلاً، كان روبرت بويل قادراً على تطبيق هذا المبدأ في صياغة قانونه حول أثر الضغط على حجم الغاز: عند الإبقاء على ثبات درجة الحرارة، يتغير حجم الغاز عكسياً مع الضغط الواقع عليه. وبشكل مماثل صاغ جاك تشارلس قانوناً يعالج أثر درجة الحرارة: عند الإبقاء على ثبات الضغط يتغير حجم الغاز طرداً مع درجة الحرارة. ونظراً لأن البحث التربوي يخص البشر فإن هناك دوماً، العديد من المتغيرات فيه. لذا، فإن محاولة تحويل المشكلات التربوية إلى عملية المتغير الواحد، ليست غير واقعية فحسب، بل حتى، ربما مستحيلة. ولحسن الحظ فإن مثل هذا الضبط الصارم ليس أساسياً بصورة مطلقة، وذلك لأن العديد من الجوانب التي تختلف فيها المواقف ليس لها صلة بهدف الدراسة، ولذا يمكن إهمالها. ويكفي أن نطبق قانون المتغير المستقل الواحد المتميز. فمثلاً، في دراسة الآثار المختلفة لطريقتين في تدريس الحساب، قد يرغب المرء في أن يكون لديه مجموعتان من الأطفال المتطابقين في كل جانب عدا الطريقة التي يتم فيها تدريسهم الحساب. ونظراً لأنه من المستحيل أن يكون لدينا مجموعتان متطابقتان من الأطفال بصورة مطلقة فإن المجرب يسعى لتكوين مجموعتين متشابهتين قدر الإمكان فيما يخص المتغيرات الأخرى ذات الصلة بتحصيل الحساب كالقدرة القرائية، والذكاء العام، والدافعية، وما أشبه.

وهمل متغيرات أخرى لا يحتمل أبدا أن تكون لها صلة بالحساب مثل القدرة الرياضية، أو الطول، أو لون الشعر. ورغم أن قانون المتغير الواحد لا يمكن اتباعه بشكل مطلق، فإن التجربة تحاول تقريبه قدر الإمكان في جميع المتغيرات ذات الصلة. وعليه فإنا نحتاج في دراسات التربية التجريبية إلى إجراءات تسمح لنا بمقارنة المجموعات على أساس المتغيرات المتميزة. وقد تم ابتكار عدد من طرق الضبط لجعل مثل هذه المقارنات ممكنة.

دعنا نفترض أننا نرغب في اختبار فرضية تقول إن الأطفال الذين تم تدريسهم حسب الطريقة الاستقرائية (المجموعة أ) يحرزون تقدماً أكبر في تعلم المفاهيم العلمية من الأطفال الذين درسوا حسب الطريقة الاستنتاجية (المجموعة ب). وبعبارة أخرى، نحن نرغب في دراسة الصلة بين طريقة التدريس (المتغير المستقل) وتعلم المفاهيم العلمية (المتغير التابع). ولكي نستخلص نتيجة بخصوص علاقة المتغير المستقل والمتغير التابع يجب أن نضبط التأثيرات لأية متغيرات "دخيلة". فالمتغير الدخيل هو متغير غير مرتبط بهدف الدراسة بل قد يؤثر على المتغير التابع. أما الضبط فهو مصطلح يستخدم لتفسير إجراءات القائم بالتجربة لإقصاء التأثيرات الفارقة لجميع المتغيرات الدخيلة بالنسبة لهدف الدراسة. وإذا كان المتغير معروفاً بكونه غير مرتبط بالمتغير التابع، فلا يمكنه عندئذ أن يؤثر عليه ولا نكون بحاجة إلى ضبط تأثيراته. وحيث لا توجد هناك صلة معروفة كالتي بين حجم حذاء طفل في سن العاشرة وقدرته على تعلم المفاهيم العلمية فإنه لا توجد هناك حاجة لدى الباحث لضبط المتغير الدخيل (في هذه الحالة حجم الحذاء).

وفي التجربة المذكورة أعلاه، يعد الذكاء عاملاً يؤثر بشكل أكيد على تعلم المفاهيم العلمية، وعليه، يعتبر متغيراً دخيلاً ذا صلة ويجب ضبطه. وبخلاف ذلك، إذا كان الأطفال في المجموعة (أ) أكثر ذكاءً من الأطفال في المجموعة (ب) فإن التحسن الأكبر في التعلم من قبل المجموعة (أ) يمكن أن يعزى إلى الذكاء، وعليه لا يمكننا أن نقيّم بشكل مناسب تأثيرات طريقة التدريس على التعلم. بعبارة أخرى، فإن الذكاء قد شوّش الصلة بين المتغيرات التي نهتم بها. ومصطلح "إرباك / تشويش" يشير إلى "الخلط" بين المتغيرات الدخيلة على مشكلة البحث والمتغيرات المستقلة للدراسة البحثية بطريقة لا يمكن فيها فصل آثارها، ولا يمكن القول بشكل واضح ما إذا كانت الصلة القائمة (1) بين متغير مستقل والمتغير التابع في الدراسة، (2) بين متغيرات دخيلة والمتغير التابع أو (3) من (1) و (2) معا. ويمكن استبعاد التشويش عن طريق ضبط تأثير المتغيرات الدخيلة ذات الصلة.

المعالجة العملية / التفعيل Manipulation

يشير معالجة / تفعيل المتغير المستقل إلى عملية مدروسة يقوم بها المجرب. وفي البحوث التربوية وغيرها من العلوم السلوكية تنطوي معالجة / تفعيل المتغير المستقل على إقامة مجموعات "معالجة" أو ظروف مختلفة. وتدعى ظروف المعالجة المختلفة التي تجرى على أفراد التجربة "مستويات" المتغير المستقل. وتمثل المستويات اثنين أو أكثر من قيم المتغير المستقل وقد تكون فروقات في

الدرجة، أو فروقات في النوع، استناداً إلى طبيعة المعالجة. وينطوي المتغير المستقل الكمي على تفعيل في مدى أو مقدار المتغير المستقل، ولذا تختلف المستويات في الدرجة أو المقدار. ويكون لدى المتغير المستقل النوعي، من ناحية أخرى، مستويات تختلف في النوع. مثلاً، افترض أن صاحب تجربة مهتم بتأثير مثير على تعلم طلبة الكلية مقاطع كلمات جوفاء. فالباحث سيبدأ بتحديد المثير الذي سيستخدم والمقدار الذي سيطبق. فإذا كان الباحث مهتماً بأثر مقدار المثير على التعلم، فإنه ربما سيبدأ بثلاثة مستويات للمتغير المستقل: مقدار عال، مقدار وسط ومقدار متدن. أو يمكن للباحث أن يقارن تأثير مثير واحد مع مثير آخر أو مع لا شيء على الإطلاق. وفي هذه الحالة تختلف المستويات في النوع.

وأمثلة أخرى للمتغيرات المستقلة حيث قد تختلف المستويات في المقدار هي الحرمان من النوم أو الأموال المستخدمة كحوافز. أما المتغيرات المستقلة النوعية كطرق التدريس المختلفة (محاضرة مقابل نقاش) أو تعليمات مختلفة مقدمة إلى الأفراد سيكون لها مستويات تختلف في نوعها. إن عدد مستويات المتغير المستقل ستساوي عدد ظروف المعالجة التي أقامها المجرب.

وقد يعالج / يفعّل الباحث متغيراً مستقلاً واحداً أو عدداً من المتغيرات في الوقت ذاته.

الرصد / الملاحظة Observation

نهتم في إجراء التجارب بأثر تفعيل متغير مستقل على متغير استجابة. وتتم حالات الرصد على بعض خواص سلوك الأفراد المستخدمين في البحث. وحالات الرصد هذه التي هي كمية في طبيعتها قدر الإمكان، تكون المتغير التابع.

إن المتغير التابع في البحث التربوي هو في الغالب تحصيل من نوع ما كالتعلم مثلاً. ونهتم غالباً بتفسير التحصيل أو التنبؤ به. لاحظ أنه ليس بوسعنا قياس التعلم مباشرة. فبوسعنا فقط تقدير التعلم من خلال قياسات كدرجات اختبار. وعليه فإنا نقول بدقة أن المتغير التابع هو عبارة عن درجات أو ملاحظات وليس تحصيلاً بذاته.

المقارنة التجريبية EXPERIMENTAL COMPARISON

تبدأ تجربة معينة بفرضية تجريبية وهي تعد تنبؤاً بأن المعالجة سيكون لها تأثير معين. وتعبر الفرضية عن التوقعات مثل الاستنتاجات الناجمة عن التغيرات التي سيتم إدخالها – أي إن مجموعات المعالجة واللا معالجة سوف تختلف بسبب تأثيرات المعالجة. ويتم التخطيط للتجربة وتنفيذها بغية جمع الأدلة ذات الصلة بالفرضية المذكورة. وبالنسبة لأبسط تجربة، هناك حاجة لمجموعتين من الأفراد: "المجموعة التجريبية" و "المجموعة الضابطة". إن التعريفات الأصلية قد حددت المجموعة التجريبية على أنها المجموعة التي تحظى بمعالجة خاصة، بينما لا تحظى المجموعة الضابطة بأية

معالجة. إن استخدام المجموعة الضابطة يُعين الباحث على تخفيض عدد من التفسيرات البديلة لتأثير المعالجة. مثلاً قد يعطي حقل، جرى تسميده، حاصلاً وفيراً بسبب الطقس المناسب أو لأسباب أخرى. فإن أعطت حقول مجاورة غير مسمّدة، حاصلاً أقل فإن تأثير السماد على الحاصل يصبح معقولاً.

وعلى أية حال، فإن الشيء المألوف على نحو أكثر، من مجرد مقارنة مجموعة المعالجة بمجموعة لم تحظ بأية معالجة (بمجموعة ضابطة حقيقية)، هو الموقف الذي نقارن فيه المجموعات التي تحظى بمعالجات مختلفة. وندعو هذه "بمجموعات المقارنة". وتدرس أغلبية التجارب التربوية الفرق في نتائج اثنتين أو أكثر من المعالجات، بدلاً من الفرق في نتائج معالجة واحدة مقابل عدم المعالجة كلياً. فمثلاً، قد يبدو الأمر خال من المعنى إذا ما قارنا تحصيل التهجئة لمجموعة تجريبية تم تدريسها حسب الطريقة (أ) مع مجموعة لم يكن لديها أي تدريس للتهجئة. فبدلاً من ذلك، تجرى مقارنة بين المجموعات التي حظيت بمعالجات الطريقة (أ) والطريقة (ب). إن مقارنة المجموعات التي حظيت بمعالجات مختلفة يؤمن الضبط نفسه لتفسيرات بديلة مثلما تقوم مقارنة المجموعات المعالجة وغير المعالجة. ولغرض تبسيط النقاشات اللاحقة سوف نستخدم مصطلح "المجموعة الضابطة" للإشارة إلى أي من المجموعات التي لم تحظ بمعالجة أو للمجموعات ذات المعالجات البديلة. وتعد المقارنات مهمة في البحث العلمي. فمقارنة مجموعة تحظى بمعالجة مع مجموعة لا تحظى بمعالجة أو بمجموعة مكافئة أو مجموعات تحظى بمعالجة بديلة تجعل من الممكن لدينا أن نستمد استنتاجات راسخة من النتائج.

ويجب أن تكون المجموعات التجريبية والضابطة متكافئة في جميع المتغيرات التي قد تؤثر على المتغير التابع، فهي تختلف فقط في تعرضها للمتغير المستقل. وبعد أن يفرض القائم بالتجربة الشروط المختلفة على الأفراد فإن كل فرد يقاس حسب المتغير التابع.

إن التقييم يلي القياس. فهل هناك فرق بين المجموعتين؟ هل يختلف تأثير معالجة (أ) عما هو عليه الحال في معالجة (ب)؟ إن هذا السؤال يفرض ويتطلب مقارنة قياسات المتغير التابع في المجموعة الأولى مع قياسات المتغير التابع في المجموعة الأخرى. ولابد أن تبين المقارنة للمجرب ما إذا كانت الاختلافات في المتغير التابع مرتبطة بالاختلافات في المتغير المستقل كما يمثلها الشرطان (أ) و (ب).

التصميم التجريبي EXPERIMENTAL DESIGN

يشير التصميم التجريبي إلى الإطار الفكري الذي تجرى التجربة ضمنه. إنه خطة الباحث لتنفيذ التجربة. وتختلف التصميمات التجريبية بطريقتين رئيسيتين. أولاً، تختلف التصميمات في عدد المتغيرات المستقلة التي يتم تفعيلها. ولبعض التصميمات التجريبية متغير مستقل واحد، بينما لتصميمات أخرى اثنان أو أكثر. ثانياً، تختلف التصميمات في طريقة تخصيص الأفراد إلى

معالجات مختلفة. ففي التجربة الحقيقية يتم تخصيص الأفراد في المجموعات بشكل عشوائي، وفي حالات أخرى يستخدم التصميم مجموعات موجودة قبلاً أو قد يحظى كل فرد بجميع المعالجات.

ويخدم التصميم التجريبي وظيفتين: 1) أنه يقيم شروطاً للمقارنات المطلوبة لاختبار فرضيات التجربة و 2) أنه يتيح للقائم بالتجربة من خلال تحليل إحصائي للبيانات، إعطاء تفسير منطقي لنتائج الدراسة. فإن تعيّن على تصميم معين إن ينجز هذه الوظائف، فيجب على المجرب أن يتذكر لدى اختياره لمثل هذا التصميم بعض المعايير العامة. إن وصف الأنماط المختلفة للتصميم التجريبي يتبع مناقشة هذه المعايير العامة.

فأهم معيار هو ضرورة أن يكون "التصميم مناسباً" لاختبار فرضيات الدراسة المعينة. وليس مؤشر التجربة هو التعقيد أو البساطة بل في كونها مناسبة. فالتصميم الذي سيقوم بالعمل الذي يجب أن يقوم به، هو التصميم الصحيح. وهكذا فأول مهمة للمجرب هو اختيار التصميم الذي يرتب الظروف التجريبية على أفضل ما يكون لتلبية احتياجات المسألة التجريبية.

أما المعيار الثاني فهو ضرورة أن يوفر التصميم "الضبط الكافي" بحيث يتسنى تقييم تأثيرات المتغير المستقل بشكل لا لبس فيه، قدر الإمكان. وما لم يضبط التصميم، المتغيرات الدخيلة، لن يستطع المرء مطلقاً، أن يكون واثقاً من العلاقة بين متغيرات الدراسة. وكما سترى فإن العشوائية هي أفضل طريقة لتحقيق الضبط الضروري. وعليه فإن أفضل نصيحة هي اختيار تصميم يستثمر العشوائية بأكثر ما يمكن من الجوانب.

إن إحدى المساهمات المهمة في تقييم التصميمات البحثية، أجراها (Campbell & Stanley, 1966) اللذان أشارا إلى أن هناك معيارين للتصميمات البحثية: "الصدق الداخلي" و "الصدق الخارجي" حيث نناقشهما فيما يلي.

الصدق الداخلي للتصميمات البحثية

Internal Validity of Research Designs

يشير الصدق الداخلي إلى المدى الذي تكون فيه التغيرات الملاحظة في المتغير التابع قد سببتها، في الحقيقة، المتغيرات المستقلة في موقف تجريبي معين. ويتعلق الصدق الداخلي بأسئلة مثل: هل سببت المعالجة التجريبية فعلاً، تغيراً في المتغير التابع؟ وهل أحدث المتغير المستقل، فعلاً، اختلافاً؟ إن هذه الأسئلة عن الصدق الداخلي، لا يمكن للمجرب الإجابة عليها بشكل إيجابي ما لم يوفر التصميم ضبطاً كاف للمتغيرات الدخيلة. فإذا وفر التصميم ضبطاً للمتغيرات، كان بوسع المرء إقصاء التفسيرات البديلة للنتيجة المرصودة، وتفسيرها على أنها توضح علاقة جوهرية بين المتغيرات. ويعد الصدق الداخلي، مشكلة ضبط، أساساً. وتصميم الضوابط المناسبة هو قضية إيجاد طرق لإقصاء المتغيرات الدخيلة – أي المتغيرات التي يمكن أن تؤدي إلى تفسيرات بديلة. فكل ما يساهم في ضبط التصميم، يساهم في صدقه الداخلي.

لقد حدد كامبل وستانلي ثمانية متغيرات دخيلة تمثل بشكل متكرر تهديدات للصدق الداخلي لتصميم البحث. ويجب ضبط هذه المتغيرات وإلا أدت إلى أثر يمكن أن يفهم خطأ على أنه أثر المعالجة التجريبية.

1- التاريخ: هناك أحداث أو ظروف معينة عدا المعالجة التجريبية، قد تحدث بين القياسين الأول والثاني للأفراد وذلك لإحداث التغييرات في المتغير التابع. ويشار إلى مثل هذه الأحداث على أنها أثر التاريخ. فمثلاً، خلال تجربة لقياس فاعلية وحدة معينة حول كيفية عمل سوق البورصة، هبط السوق بشكل حاد وتركز وسائل الإعلام اهتماماً كبيرا عليه. ولا يمكن للباحث أن يحدد ما إذا كانت المعرفة الأكبر للطلبة عند نهاية الوحدة، تعود للوحدة أو لتعرض الطلبة إلى التغطية الإعلامية. فتأثيرات الوحدة وتعرض الطلبة للتغطية الإعلامية متشابكة، ويستحيل معرفة مقدار تعلم الطلبة بسبب الوحدة ومقدار تعلمهم بسبب التاريخ كمسألة صدق داخلي.

2- النضج: إن العمليات التي تعمل داخل الأفراد بفعل مرور الزمن قد تنتج آثاراً يمكن أن نعزوها خطأً إلى المتغير التجريبي. وتمثل هذه العمليات تهديد النضج للصدق. فقد يعمل الأفراد بصورة مختلفة على قياس المتغير التابع لأنهم ببساطة أكبر عمراً أو أشد جوعاً، أو أكثر تعباً، أو أقل دافعية مما كانوا عليه في وقت القياسات الأولى. فمثلاً، قد يصعب تقدير آثار المعالجات لمشكلات النطق بين الأطفال قبل سن الدراسة، لأن الأطفال الصغار غالباً ما يتخلصون من مثل هذه المشكلات بصورة طبيعية.

3- إجراء الاختبار القبلي: قد يؤثر التعرض للاختبار القبلي قد يؤثر على أداء الأفراد في الاختبار الثاني، بغض النظر عن المعالجة التجريبية. وقد يتعلم الأفراد المادة من الاختبار القبلي. فقد يؤدون بشكل أفضل في الامتحان البعدي لأنهم أصبحوا يألفون صيغة الاختبار والبيئة الاختبارية، أو طوروا استراتيجية للأداء الحسن في الاختبارات، أو أصبحوا أقل قلقاً في المرة الثانية. فأثار الممارسة هذه قد تسبب درجات اختبار بعدي أعلى عندما لا يكون للمتغير المستقل أي تأثير على الإطلاق. وفي حالة استبيانات الاتجاهات والشخصية، فإن أخذ الاختبار قد يحفز على التفكير بالأسئلة والقضايا التي تثار في الاختبار القبلي، فيعطون استجابات مختلفة للاختبار البعدي.

4- أدوات القياس: إن التغيرات في أدوات القياس، أو في أصحاب الدرجات، أو في الراصدين / الملاحظين قد يؤدي إلى تغيرات في القياسات الناتجة. فإذا كان الاختبار البعدي أكثر صعوبة أو إذا تم استخدام راصدين مختلفين لقياسات قبلية وبعدية، فإن هذه العوامل قد تفسر الفروقات المرصودة في الدرجتين. وإذا عرف الراصدون الذين يسجلون المتغير التابع، الأفراد الذين تلقوا المعالجة والذين لم يتلقوها، فإنهم قد يبالغون في تقدير أداء المعالجة، بشكل غير واعٍ، ويقللون من تقديرها بالنسبة للمجموعة الضابطة.

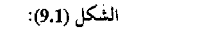
5- الارتداد الإحصائي(*): إذا تم اختيار المجموعات على أساس الدرجات القصوى فقد يعمل الارتداد الإحصائي على إنتاج أثر يمكن أن يفسر خطأ على أنه أثر تجريبي. ويشير أثر الارتداد هذا، إلى ميل الدرجات المتطرفة نحو الارتداد أو التحرك باتجاه الوسط الحسابي للقياسات اللاحقة.

فمثلاً، دعنا نفترض أن أصحاب الدرجات الدنيا من الصف الرابع في اختبار الكفاءة بالإنجليزية قد تم اختيارهم لبرنامج تجريبي خاص بالإنجليزية. إن الوسط الحسابي لهذه المجموعة سوف يميل إلى الصعود / الارتداد نحو وسط المجتمع الإحصائي في الاختبار الثاني سواء تم تطبيق المعالجة التجريبية أم لا. وبصورة مماثلة يميل أصحاب الدرجات العليا الأولى إلى الهبوط / الارتداد نحو وسط المجتمع الإحصائي في الاختبار الثاني.

دعنا نوضح الارتداد مع بيان التشتت (الشكل 9.1) الذي يوضح النمط الذي نحصل عليه إذا كان ارتباط درجات اختبار القراءة للصف الرابع ودرجات اختبار القراءة للصف الخامس هو ($r = 0.7$). فكل نقطة تمثل كلا من الدرجتين z لفرد ما. وإذا اخترنا أفراداً لهم درجة z معينة (x) في اختبار القراءة للصف الرابع وتفحصنا درجاتهم في اختبار القراءة للصف الخامس، لوجدنا أن لدى القليل منهم الدرجة z ذاتها في كلا الاختبارين. وإذا كانت الدرجة z لهذه المجموعة الجزئية من الصف الرابع، فوق الوسط الحسابي، لوجدنا أن القليل من الطلبة يحرزون درجة يسجلون أعلى بكثير من الوسط في الصف الخامس بما يفوق ما حصلوا عليه في الصف الرابع، لكن لغالبيتهم درجات z الأقرب إلى الوسط، بل حتى إن بعضهم يكونون دون الوسط الحسابي.

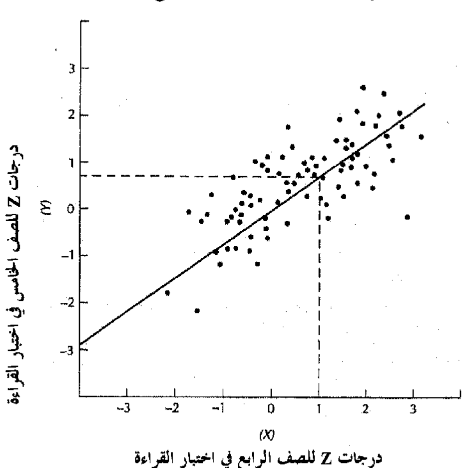
بالنسبة إلى المجموعة الجزئية التي لها درجة z بانحراف معياري واحد، فوق الوسط على (x)، يكون فإن الوسط على (y) هو (+0.7). أما بالنسبة إلى المجموعة الجزئية ذات درجة z بانحرافين معياريين دون الوسط على (x) فإن الوسط الحسابي على y هو (-1.4)، والمجموعة ذات نقطة z (+2) على (x) لها وسط درجة z هو (+1.4) على (y) وهكذا. إن الخط المائل في بيان التشتت يربط هذه الأوساط على (y) بدرجات مختلفة على (x). لاحظ أن لكل مجموعة ذات درجات فوق الوسط الحسابي في اختبار القراءة للصف الرابع تكون درجة الوسط على (y) أدنى في اختبار القراءة للصف الخامس. وبالنسبة لكل مجموعة ذات درجات أدنى من الوسط في اختبار القراءة للصف الرابع فإن وسط الدرجات هو "أعلى" في اختبار القراءة للصف الخامس. بعبارة أخرى، إن متوسط الدرجات لكل مجموعة يتحرك أقرب إلى الوسط الحسابي – وهي نتيجة الارتداد.

(*) ارتداد (Regression): مصطلح كان أول من استخدامه Galton و Pearson لوصف ظاهرة وراثية حول علاقة أطوال الآباء والأبناء. لقد لاحظا أن طول الأبناء من أبوين طويلين يكون أقل من طولي لآباء، وبالعكس إذا كان الأبوان قصيرين كان الأبناء أطول من الأبوين، بمعنى أن طول الأبناء يرتد نحو متوسط طول المجتمع. وهناك من يترجمه: الانحدار أو التراجع أو الانكفاء. ونرى أن "الارتداد" قريب من الفهم العام، فهو أقرب للتعلم. (المراجع)



الشكل (9.1):

بيان التشتت لدرجات اختبار القراءة للصفين الرابع والخامس



يحدث الارتداد بصورة حتمية عندما يكون الارتباط بين متغيرين أقل من التام. وحيث لا نوجد متغيرات ذات أهمية في التربية، مترابطة بشكل تام من الناحية العملية، فعلينا دائماً أن ندرك أثر الارتداد في تصميم تجاربنا. ويتضح جانب الارتداد في المثل القائل: "عندما تكون في الحضيض فليس أمامك خيار سوى أن تصعد" و "عندما تكون في القمة فليس أمامك خيار سوى أن تهبط".

مثلاً، تعد مدرّسة برنامج علاج في دروس الرياضيات، وتخصص لهذا البرنامج طلبة المدرسة المتوسطة ممن كانت درجاتهم سنتين أو أقل دون مستوى فرقتهم، حسب اختبار مقنّن للرياضيات. وبعد فصل دراسي يعطى للطلبة شكل مكافئ للاختبار المقنن ويحرز غالبيتهم درجات قريبة من مستوى فرقتهم، بما يبدوا أنهم استفادوا من البرنامج. وقبل أن نعزو المكاسب إلى هذا البرنامج، علينا أن نتذكر أنه حين يكون المرء في الحضيض، فليس لديه خيار سوى الصعود. وتشمل الدرجات الأدنى في الاختبار القبلي، الطلبة الذين تدنت درجاتهم بسبب سوء

الصحة المؤقت، أو المشكلات العاطفية، أو عدم الانتباه، أو مشكلات أخرى. وكذلك، سيشمل أولئك الذين تدنت درجاتهم بسبب خطأ القياس. ويتوقع من غالبية هذه الحالات أن تحصل على درجات أفضل في الاختبار اللاحق. وبصورة معكوسة، إذا خصص إلى الذين حصلوا على درجات عالية في مقياس ما، معالجة خاصة مثل برنامج إثرائي فإن قياس المعالجة البعدية، قد يجعل الأمر يبدو وكأن للمعالجة أثراً ضاراً. فليس لمن في القمة خيار، سوى الهبوط إلى الأسفل.

6- **الاختيار التمييزي / الفارقي للأفراد:** قد تكون هناك فروقات مهمة بين المجموعات حتى قبل تطبيق المعالجة التجريبية. مثلاً، إذا كانت المجموعة التجريبية في تجربة للتعلم أكثر ذكاءً من المجموعة الضابطة، فمن المتوقع أن تؤدي الأولى بشكل أفضل حسب قياس المتغير التابع، حتى وإن لم تحظ هذه المجموعة بمعالجة تجريبية.

7- **التسرب التجريبي:** يحدث خطر تسرب حين تكون هناك خسارة تمييزية / فارقية للمستجيبين من مجموعات المقارنة. فإن تسرب فرد معين من المجموعة خلال دورة التجربة، فإن هذه الخسارة الفارقية قد تؤثر على نتيجة الدراسة. وإذا تسرب، مثلاً، عدد من أصحاب الدرجات المتدنية في الامتحان القبلي بشكل تدريجي من المجموعة التجريبية، فسوف يكون لهذه المجموعة وسط أداء عال في القياس النهائي، ولا يكون ذلك بسبب المعالجة التجريبية، بل بسبب غياب الأفراد ذوي الدرجات المتدنية عند إجراء الاختبار البعدي. فمثلاً، أرادت مستشارة كنيسة أن تقارن فاعلية اثنين من الإجراءات الاستشارية المختلفة في أمور الزواج. وقد أجرت قياساً لما قبل المعالجة حول تكيف الأزواج المشمولين بالبرنامج، ثم وزعت عشوائياً كل زوج حسب الإجراء (أ) أو (ب). ومع الإجراء (أ) الذي ضم تفحص الذات(*) والمواجهة، استمر فقط (11) زوجاً من (20) في البرامج وكانوا حاضرين لقياس ما بعد المعالجة. ومع الإجراء (ب) الذي كان أكثر راحة وأقل إلحاحاً فإن 18 من (20) زوجاً استمروا في البرنامج، وكانوا حاضرين لقياس ما بعد المعاجلة. لقد كان وسط ما بعد المعالجة، للإجراء (أ) أعلى من وسط مجموعة الأجراء (ب). وقد يكون ذلك دليلاً على إن الإجراء (أ) هو أكثر فاعلية، لكنه قد يكون دليلاً أيضاً على أن الأزواج الذين كانت دافعيتهم عالية لتحسين زواجهم استمروا في البرنامج، بينما شملت درجات ما بعد المعالجة للإجراء (ب) الأزواج الأقل دافعية.

8- **تفاعل الاختيار والنضج:** قد تتفاعل بعض هذه الأخطار. وقد يحدث مثل هذا التفاعل في تصميم شبه تجريبي حيث لا يتم اختيار المجموعات التجريبية والضابطة بشكل عشوائي، بل كمجموعات تامة موجودة بشكل مسبق مثل الصفوف المدرسية. ورغم ذلك قد يشير اختبار قبلي إلى أن المجموعات متكافئة في بداية التجربة، وقد يتصادف أن تكون المجموعة

(*) تفحص الذات (soul Searching): مصطلح يشير إلى قيام المرء بتفحص ضميره ومبادئه، ومدى انسجامها مع أفعاله (المراجع).

التجريبية ذات معدل نضوج أعلى من المجموعة الضابطة، وإن الزيادة في معدل النضوج هو الذي يعزى إلى الأثر المرصود. فعندما يتم انتخاب الطلبة الأسرع نضجاً في المجموعة التجريبية، فإن أثر تفاعل الاختيار والنضج قد يعزي خطأً للمتغير التجريبي.

فمثلاً، إذا اختارت مجموعة من طلبة السنة الأولى بالكلية دراسة مقرر متقدم في الإنجليزية، فقد تنمو مفرداقم أكثر من مجموعة في السنة ذاتها تدرس مقرراً اعتياديا في الإنجليزية، وذلك لأن معدل نمو المفردات كان أعلى قبل، وخلال السنة الأولى. إن مقارنة الفروقات بين الاختبارين القبلي والبعدي للمقررين المتقدم والاعتيادي قد تدفع بأحدهم إلى الاستنتاج بصورة خاطئة بأن المتغير المستقل كان مسؤولاً عن الفرق في الكسب الذي يعزى فقط إلى تفاعل الاختيار والنضج. وتبرز هذه المشكلة تكراراً عندما يقارن المتطوعون مع غير المتطوعين. فالمتطوعون قد يكونون أكثر دافعية لتحقيق فائدة حسب المتغير التابع، أعلى مما هو عليه الحال مع غير المتطوعين، وقد يعزى هذا الفرق في الفائدة / الكسب خطأً إلى المتغير المستقل – كما يمكن أن يحدث حتى عندما تكون المجموعات متكافئة في الاختبار القبلي.

وفي الوقت الذي أدرج كامبل وستانلي فقط ثمانية أخطار قمدد الصدق الداخلي، فقد أشار بعض المؤلفين إلى إن هناك على الأقل اثنين (أو ثلاثة) من الأخطار الأخرى التي ينبغي دراستها.

9- التنفيذ: أحياناً، يهدد التنفيذ الفعلي للتجربة الصدق الداخلي، كالحالة التي تعطى فيها المجموعة التجريبية، سهواً، أفضلية غير مخطط لها، على المجموعة الضابطة. فمثلاً، في تجربة لمقارنة فاعلية اثنتين من طرق التدريس قد يخصص المدرس الأكثر اقتداراً للمجموعة التجريبية.

وقد يعمل هذا الخطر إذا كان للقائم بالتجربة تحيزاً شخصياً لصالح طريقة على أخرى. وقد تنتقل هذه المعاملات الخاصة والتوقعات من القائم بالتجربة، بصورة غير واعية، إلى الأفراد بطريقة تؤثر على سلوكهم. وتدعى هذه الظاهرة "أثر / ظاهرة تحيز المجرب". فمثلاً، إفترض أن المجرب يبحث في فاعلية أسلوب تدريس جديد قد قام بتطويره ويعتقد أنه ممتاز. فإذا كان المجرب هذا في موقع يؤثر فيه على التجربة، فإنه قد يتصرف عن دون قصد، بطريقة تحسن أداء المجموعة التجريبية وليس المجموعة الضابطة. والباحث غير المتحيز لن يحصل بالضرورة على الاستنتاجات ذاقما. وقد تقصى (Rosenthal, 1985) ظاهرة تحيز المجرب بشكل واسع. فقد أوضح أن الفئران التي حددت بوصفها "ذكية" فاقت في أدائها ما يدعى بالفئران "الغبية" وذلك بإقناع المجربين ببساطة بأن هذه التسميات كانت صحيحة. وكانت الدرجات العالية للفئران "الذكية" تعود إلى تحيز المجرب وليس بسبب أية فروقات حقيقية بين الفئران. وينبغي، على أية حال، التأكيد على إن أثر التحيز هذا، لا ينتج عن أي فعل مدروس من جانب المجرب لتغيير البيانات أو لجعل التجربة تظهر بطريقة معينة.

إن أفضل طريقة للتعامل مع أثر / ظاهرة تحيز المجرب هي بجعل الأفراد المدربين الآخرين (وليس الباحث نفسه) يعملون بشكل مباشر مع المشاركين في الدراسة. وينبغي عليه أن يحجم عن إبلاغ الذين يجرون التجربة بأن معالجة تجريبية معينة تعد أفضل من الأخرى.

10- اتجاهات الأفراد: يمكن أن تكون اتجاهات الأفراد تهديداً للصدق الداخلي. فمثلاً، في دراسة أثار مستويات الإضاءة المختلفة على إنتاجية العمال في مصنع هوثورن في ولاية الينوي، يعود لشركة ويسترن أليكتريك، لاحظ الباحثون أن الزيادة والنقصان في شدة الإضاءة نجم عن كليهما إنتاجية متزايدة. وعلى أية حال، فقد استنتج الباحثون بأن الاهتمام الذي أعطي للعمال ومعرفتهم بأنهم كانوا يساهمون في تجربة - وليس بالتغيرات في الإضاءة - كانت العوامل الرئيسية التي أدت إلى مكاسب الإنتاج. إن هذا الميل لدى الأفراد في تغيير سلوكهم فقط لكونهم يساهمون في تجربة، قد أشير إليه لاحقاً بكونه "أثر / ظاهرة هوثورن". ويمكن أن يشكل هذا مشكلة في البحوث التربوية التي تقارن طرق التدريس الجديدة مع الطرق التقليدية.

أما عكس أثر هوثورن فهو "اثر / ظاهرة جون هنري". ويشير هذا إلى ميل أفراد المجموعة الضابطة، ممن يعرفون أنهم في تجربة كي يمارسوا جهداً إضافياً ويؤدوا بناء على ذلك عملاً فوق معدلهم الاعتيادي أو المتوقع. وقد يتصورون أنهم في حالة تنافس مع المجموعة التجريبية ويريدون أن يؤدوا العمل مثلهم أو أفضل. وهكذا فإن الاختلاف (أو الافتقار إلى الاختلاف) بين المجموعات قد يعزى إلى الدافعية المتزايدة لدى أفراد المجموعة الضابطة وليس إلى المعالجة التجريبية. ويحتمل أن يحدث هذا التأثير في بحث صفي يقارن فيه أسلوب تعليم جديد بطريقة تقليدية قد يحل الأول مكانها. وقد يرغب الطلبة في الصف التقليدي أن يوضحوا أن بوسعهم الأداء مثل الطلبة الذين يتعلمون حسب الطريقة الجديدة. وفضلاً عن ذلك، قد يرغب المدرس في صف المجموعة الضابطة أن يوضح بأن الطريقة القديمة هي جيدة أيضاً، فتكون لهم دافعية لبذل جهد خاص. مثلاً، أوضحت البحوث الأولى حول التدريس بواسطة التلفزيون، أن المدرسين في الصف الاعتيادي (المجموعة الضابطة) قد بذلوا جهداً خاصاً كي يضاهي أداء تلاميذهم أو يفوق أداء الطلبة الذين تلقوا تدريساً تلفزيونياً (المجموعة التجريبية). وربما شعر المدرسون في الصف التقليدي بخطر هذا الابتكار وأرادوا أن يبينوا أن بوسعهم التدريس بشكل أفضل من التلفزيون.

ضبط الفروقات بين الأفراد Controlling Intersubject Differences

ينبغي أن تصمم التجربة كي تتجنب، أو على الأقل تقلص، أثر التهديدات للصدق الداخلي. فيجب أن توجه الجهود الأولى للمرء نحو ضبط أية فروقات موجودة أصلاً بين الأفراد المشاركين في التجربة. وفقط، بهذه الطريقة يمكن للمرء أن يكون واثقاً إلى حد ما بأن أية فروقات بعد التجربة يمكن أن تعزى إلى ظروف التجربة أكثر مما تعزى إلى فروقات الأفراد الموجودة أصلاً. وهناك ستة إجراءات أساسية تستخدم بشكل مألوف لزيادة التكافؤ بين

المجموعات المعرّضة للأوضاع التجريبية المختلفة وهي 1) التخصيص العشوائي، 2) التناظر العشوائي، 3) الاختيار المتجانس، 4) إدماج متغيرات في التصميم، 5) تحليل التباين الاقتراني، 6) استخدام الأفراد كضابطين لأنفسهم.

التخصيص العشوائي: لندرس مهمة المجرب. هناك مجموعة متيسرة من الأفراد الذين يجب تقسيمهم من أجل التجربة إلى مجموعتين، تعالجان بصورة مختلفة ثم تتم مقارنتهما مع بعضيهما. وبتخصيص الأفراد إلى المجموعتين من أجل التجربة، يحتاج المجرب إلى نظام يعمل بشكل مستقل عن الحكم الشخصي، وخواص الأفراد أنفسهم. مثلاً، يجب أن لا يخصص أصحاب الدرجات العالية المعروفين للمجموعة (أ) وأصحاب الدرجات المتدنية للمجموعة (ب). فالنظام الذي يلبي هذا الشرط هو "التخصيص العشوائي". والتخصيص العشوائي هو تخصيص الأفراد إلى مجموعات بطريقة يكون فيها، لكل تعيين مفروض، احتمال متساو لكل فرد في المجتمع الإحصائي، في التخصيص لأي من المجموعات. ومصطلح (العشوائية) غالباً ما يستخدم كمرادف للتخصيص العشوائي.

إن التخصيص العشوائي لا يماثل الاختيار العشوائي، فالاختيار العشوائي (أنظر الفصل 6) هو استخدام إجراء الصدفة لاختيار عينة من المجتمع. والتخصيص العشوائي هو استخدام إجراء الصدفة لتخصيص الأفراد في المعالجات. مثلاً، لدى الآنسة براون (1000) فرد متوفرين لتجربة، ولكن التسهيلات المتاحة تكفي فقط لضم (100) فرد في تجربتها. وباختيار (100) فرد عشوائياً لضمهم في التجربة، تتجنب الآنسة بروان خلق فروقات منتظمة بين الأفراد في التجربة والمجتمع الإحصائي الذي تم اختيارهم منه. وعشوائياً بتخصيص (50) فرداً من (100) للمعالجة (أ) و (50) (ب) تتجنب الآنسة بروان فروقات ما قبل المعالجة المنتظمة بين المجموعتين.

وللحصول على مجموعات عشوائية. يمكن للباحث أن يرقم جميع الأفراد المتوفرين، ثم يسحب من جدول الأعداد العشوائية العدد المطلوب للمجموعتين. وبرمي قطعة نقدية بعدئذ يمكن أن تحدد أي المجموعتين ستحصل على المعالجة (أ) والأخرى على المعالجة (ب).

عندما يتم تخصيص الأفراد عشوائياً إلى مجموعات، فيمكن اعتبار هذه المجموعات "متكافئة إحصائياً". والتكافؤ الإحصائي لا يعني أن المجموعات متساوية بشكل مطلق، لكنه يعني إن أي فرق بين المجموعات يعود للصدفة وحدها، وليس هناك دور لتحيز الباحث، أو اختيارات الأفراد، أو لأي عامل آخر. فيحتمل لفرد ذي ذكاء عال أن يخصص للمعالجة (أ) كما للمعالجة (ب). ويصح الشيء ذاته للفرد ذي الذكاء المنخفض. أما بالنسبة لمجموع العينة فإن آثار الذكاء على المتغير التابع سوف تميل إلى التوازن أو التوزيع العشوائي. وبذات الطريقة، فإن فروقات الأفراد في الآراء السياسية، والمزاج، والدافعية، والخواص الأخرى، سوف تميل إلى أن تكون موزعة بشكل متساو تقريباً بين المجموعتين. وكلما كان لدى الباحث أفراد أكثر، كان للتخصيص العشوائي احتمال أكبر في الحصول على مجموعتين متكافئتين تقريباً.

عند استخدام التخصيص العشوائي، تعود الفروقات قبل المعالجة بين المجموعات للصدفة وحدها. وحين يكون الأمر على هذه الحال، فإنه يمكن للإحصاء الاستدلالي أن يحدد مدى احتمال، كون الفروقات قبل المعالجة، تعزى إلى الصدفة وحدها.

لاحظ أنه لا يتوقع للمتغيرات الدخيلة المعروفة أن تتوزع عشوائياً فحسب، بل أيضاً المتغيرات الدخيلة الأخرى ذات الصلة التي لا يعرفها الباحث أو يتصورها.

التناظر(*) العشوائي: إن إجراءاً بديلاً لتخصيص الأفراد إلى مجموعات، يكون عن طريق مزاوجة / تناظر الأفراد حسب كل المتغيرات الدخيلة التي يعتقد الباحث أنها تؤثر على المتغير التابع، ويتبع ذلك استخدام إجراء عشوائي ما، لتخصيص فرد كل زوج متكافئ في المعالجة (أ) وآخر في المعالجة (ب). وإذا تناظرت المجموعتان بشكل كاف، حسب هذه المتغيرات، كان هناك تأكيد معقول بأن الفروقات بعد التجربة تعزى إلى المعالجة التجريبية.

ورغم أن التناظر طريقة لتوفير الضبط الجزئي للفروقات بين الأفراد، فإنه تظل هناك العديد من الصعوبات التي قد يواجهها الفرد. وتتمثل أولى هذه الصعوبات في تحديد المتغيرات التي ستستخدم للتناظر، فمتغيرات مثل IQ، أو العمر العقلي MA، أو الوضع الاجتماعي الاقتصادي (SES)، أو العمر، أو الجنس (Gender)، أو درجة القراءة، أو درجة الاختبار القبلي، قد يتم استخدامها. فالمتغيرات التي يتناظر الأفراد، وفقا لها، يجب أن تكون ذات ارتباط قوي بالمتغير التابع، وإلا سوف يكون التناظر دون فائدة. وكقاعدة عامة نشير إلى أنه ما لم يترابط المتغير بمقدار (0.50) أو أعلى مع المتغير التابع، فلا ينبغي استخدامه لإجراء التناظر لأنه سيقوم بالقليل في زيادة دقة الدراسة. وبصورة نموذجية، نرغب بالتناظر حسب متغيرين أو أكثر يرتبط كل منها بشكل جيد مع المتغير التابع ولا ترتبط بشكل دال مع بعضها. وعلى أية حال، عندما نحاول التناظر حسب أكثر من متغيرين، يصبح من المستحيل تقريباً إيجاد أفراد يتناظرون بشكل جيد حسب هذه المتغيرات. ويفقد أفراد لعدم وجود من يناظرهم.

ويبرز سؤال آخر هنا، حول مدى قرب تناظر الأفراد حسب المتغيرات. فالتناظر بشكل وثيق، يزيد من دقة الطريقة، غير أنه يزيد أيضاً من عدد الأفراد الذين لا يمكن إيجاد من يناظرهم. وهذا، بالطبع، يقلص من حجم العينة ويدخل تحيز المعاينة إلى الدراسة.

ويجب على الباحث أن يقرر إجراء تناظر يكون عملياً في كل وضع معين. والطريقة الاعتيادية هي استخدام إجراء شخص – لشخص، الذي يبذل فيه جهد لتحديد موقع شخصين بين الأفراد الموجودين، ممن لهم درجات ضمن الحدود الموصوفة. فمثلاً، إذا كان متغير التناظر هو IQ، فإن الباحث يحدد موقع فردين هما ضمن، فارق، (5) درجات مثلاً، بينهما على مقياس

(*) التناظر (matching): مصطلح يشير إلى مزاوجة عنصرين يتفقان، إلى حد بعيد، في خواص مشتركة. وتكون المزاوجة تطابقاً، إذا اتفق العنصران في كل شيء، ويشار لها "بالتوأمة"، وهي حالة صعبة المنال. (المراجع)

IQ، ثم يخصص فرداً واحداً بشكل عشوائي للمعالجة (أ) والمناظر له للمعالجة (ب). ولن يكون، من الصعب، تناظر الأفراد على متغير IQ، فحسب. لكنه إذا كان الجنس (ذكر أو أنثى)، والطبقة الاجتماعية، متغيرين على صلة كذلك، فإنه سيصبح من الصعب جداً إيجاد أزواج متناظرة حسب مستوى المتغيرات الثلاثة. والأفراد الذين لا يمكن إيجاد تناظر لهم يعتبرون مفقودين بالنسبة للباحث.

وثمة طريقة أخرى للتناظر، تكون بوضع جميع الأفراد في سياق ترتيبي على أساس درجاتهم حسب متغير التناظر. ويتم اختيار أول فردين من قائمة سياق الترتيب (بغض النظر عن الفرق الفعلي في درجاتهم) وذلك لتكوين الزوج الأول. ثم يتم تخصيص فرد واحد من الزوج الأول بصورة عشوائية للمعالجة (أ) والآخر للمعالجة (ب). ويتم اختيار الفردين التاليين في القائمة ثم مرة أخرى يتم بتخصيص عشوائي، واحد إلى (أ) والآخر إلى (ب). ويجري الاستمرار في هذه العملية حتى يتم تخصيص جميع الأفراد. ومن السهل، نوعاً ما، إجراء التناظر حسب هذا الإجراء، إلا أنها أقل دقة من طريقة شخص - لشخص. لاحظ إن التناظر العشوائي يتطلب تناظر الأفراد أولاً، ثم تخصيصهم عشوائياً للمعالجات. والدراسة التي يقارن فيها أفراد يتبعون معالجة واحدة مع أفراد يتبعون معالجة أخرى، لا يمكن تصنيفها على أنها دراسة تجريبية. فمثل هذه الدراسات (التي نوقشت في الفصل 10) حيث يكون التناظر قائماً، إلا أن التخصيص العشوائي للمجموعات ليس قائماً، يمكن أن تقود الباحثين إلى استنتاجات خاطئة.

الاختيار المتجانس: وبوسع طريقة أخرى أن تجعل المجموعات قابلة للمقارنة بصورة معقولة حسب متغير دخيل، وذلك باختيار عينات متجانسة قدر الإمكان حسب ذلك المتغير. فإذا شك الباحث أن العمر متغير قد يؤثر على المتغير التابع، فيجرى اختيار الأطفال حسب عمر معين فقط. فباختيار أطفال السادسة، مثلاً، سيضبط الباحث أثار العمر كمتغير مستقل دخيل. وبصورة مماثلة، إذا كان محتملاً للذكاء أن يكون متغيراً يؤثر على المتغير التابع في الدراسة، فإن الأفراد سيتم اختيارهم من الأطفال ضمن مدى محدد من معامل الذكاء IQ لنقل، 100-110. وبهذا الإجراء يتم ضبط تأثير IQ. ثم يخصص الباحث بصورة عشوائية، الأفراد في مجموعات من المجتمع الإحصائي المتجانس الناتج، وبوسعه أن يكون واثقاً بأنهم قابلون للمقارنة حسب IQ. إن البدء مع مجموعات متجانسة على أساس المتغير ذي الصلة، يقصي صعوبة محاولة مقارنة الأفراد حسب ذلك المتغير.

ورغم أن الاختيار المتجانس يعد طريقة فاعلة لضبط المتغيرات الدخيلة، فإن عيبه هو تقليص المدى الذي يمكن فيه تعميم الاستنتاجات على أوضاع أخرى. فإذا كان الباحث يدرس فاعلية طريقة معينة مع مثل هذه العينة المتجانسة، لنقل، الأطفال حسب IQ متوسط، فإن النتائج لا يمكن تعميمها على أطفال حسب مدى IQ آخر. ففاعلية الطريقة مع أطفال ذوي ذكاء منخفض أو ذكاء عال جداً، سوف لا يمكن معرفتها، لذا سوف يتوجب تكرار التجربة مع أفراد من فئات IQ مختلفة.

وكما هو الحال مع التناظر، فإن التجربة الحقيقية تتطلب اختيار جميع الأفراد أولاً، ثم تخصيصهم عشوائياً للمعالجات.

إدماج متغيرات في التصميم: يمكن دمج بعض المتغيرات المرتبطة بالأفراد أنفسهم، في تصميم التجربة، وهكذا يمكن ضبطها. فمثلاً، إذا أريد ضبط الجنس (مذكر / مؤنث) في تجربة معينة، واختار المرء عدم استخدام أسلوب الاختيار المتجانس أعلاه، فإن المرء يمكن أن يضيف الجنس كمتغير مستقل آخر. وسيشمل الباحث كلاً من الذكور والإناث في الدراسة، ثم يستخدم تحليل التباين لتحديد آثار كل من الجنسين والمتغير المستقل الرئيسي على المتغير التابع. ولا تضبط هذه الطريقة المتغير الدخيل للجنس، فحسب، بل توفر، أيضاً، معلومات تأثيره على المتغير التابع وتفاعله مع المتغيرات المستقلة الأخرى.

تحليل التباين الاقتراني: يعتبر تحليل التباين الاقتراني (ANCOVA) إجراءاً إحصائياً يحسّن من دقة تصميم البحث، باستخدام متغير موجود مسبقاً ذي ارتباط بالمتغير التابع. انظر، مثلاً، في تجربة معينة لدراسة أثار طريقتين لتدريس القراءة. ففي بداية الفصل الدراسي، يجرى اختبار قبلي ويتم تخصيص نصف الأفراد عشوائياً للطريقة (أ) ونصف للطريقة (ب). فيكون المتغير المستقل، طريقة تدريس القراءة، والمتغير التابع هو الاختبار البعدي للقراءة. والاختبار القبلي للقراءة هو متغير التباين الاقتراني Covariate، أي المتغير المرتبط بالمتغير التابع، الذي يستخدم لإضافة الدقة على التصميم.

فالأفراد الذين يعتبرون قارئين جيدين، سينـــزعون للحصول على درجات جيدة في كل من الاختبارين القبلي والبعدي، بينما سينــــزع القارئون الضعاف للحصول على درجات متدنية في كلا الاختبارين، وعليه، سيكون هناك ترابط موجب بين درجات الاختبارين القبلي والبعدي. إن أسلوب تحليل التباين الاقتراني (Ancova) يزيل جزءاً من درجة الاختبار البعدي لكل فرد، أي تلك التي تعد مشاركة في درجة الاختبار القبلي.

وباستخدام هذا الأسلوب لا نأخذ بنظر الاعتبار درجة الاختبار البعدي للفرد بحد ذاتها. فنحن نحلل، بدلاً عن ذلك، الفرق بين درجات الاختبار البعدي وما نتوقع أن تكون عليه درجات الاختبار البعدي مع الأخذ بنظر الاعتبار درجة الاختبار القبلي، والارتباط بين الاختبارين القبلي والبعدي.

وباستبعاد ذلك الجزء من تباين درجة المتغير التابع المرتبطة بشكل منتظم مع تباين الاختبار القبلي، فإن دقة التجربة سوف تتحسن. ومع استبعاد ذلك الجزء من التباين في درجات الاختبار البعدي التي لا تعزى للمعالجة، فإن الفرق الذي يعزى إلى المعالجة سوف يظهر بوضوح. وباستخدام متغير التباين الاقتراني (Covariate) المرتبط بالمتغير التابع نقلص من احتمال خطأ النوع II (2).

ويمكن استخدام مقاييس أخرى، فضلاً عن درجات الاختبار القبلي كمتغيرات تباين

اقتراني. ففي مثالنا، يمكن استخدام مقاييس مثل درجات اختبار الذكاء أو معدل النقط (GPA). ولكي يكون متغير التباين الاقتراني مفيداً، ينبغي له أن يكون مترابطا (correlated) مع المتغير التابع. وفي المثال، إذ لن تكون درجات رمية كرة البيسبول أو أحجام الأحذية متغيرات تباين اقتراني مفيدة، كما يتوقع أن يكون لها ترابط، يمكن إهماله، مع الاختبار البعدي للقراءة.

استخدام الأفراد كضابطين لأنفسهم: هناك إجراء آخر ينطوي على تخصيص الأفراد أنفسهم لجميع الظروف التجريبية، والحصول على قياسات للأفراد، أولاً تحت معالجة تجريبية واحدة، ثم تحت أخرى. فمثلاً، قد تكون هناك حاجة لذات الأفراد كي يتعلموا قائمتين مختلفتين من المقاطع التي تخلو من المعنى، قائمة ذات قيمة تداع عال للأفكار وأخرى ذات قيمة تداع منخفضة. وبإيجاد الفرق في وقت التعلم بين القائمتين لكل فرد، ومتوسط الفرق في وقت التعلم لكل الأفراد، يمكن بعدئذ اختبار دلالته.

وهذه طريقة فاعلة للضبط عندما تكون عملية، إلا إن هناك ظروفاً لا يمكن فيها استخدامها. ففي بعض أنواع الدراسات، قد يجعل التعرض لظرف تجريبي واحد من المستحيل استخدام الأفراد أنفسهم تحت ظرف تجريبي آخر. فنحن لا نستطيع، مثلاً، تعليم الأطفال التقسيم على كسر، ثم نمحو ذاكرتهم، وندرسهم إياه بطريقة أخرى.

ففي التجربة السابقة، حيث استخدمت مجموعة واحدة من الأفراد للتحقق من السهولة النسبية في تعلم مقاطع الكلام الخالية من المعنى ذات التداعي العالي والمنخفض، قد يكون هناك تأثير "تعلم التعلم"، لذا فأية قائمة تظهر ثانية سيكون لها أرجحية على الأولى. وبصورة معكوسة، فإن آثار التعب أو التداخل قد تؤدي إلى أداء ضعيف في القائمة الثانية. وفي أي من الحالين لا نستطيع أن نفصل أثر المتغير المستقل وأثر الترتيب على المتغير التابع. إن استراتيجية مفيدة لهذه التجربة تكون بتقسيم الأفراد بصورة عشوائية إلى مجموعتين: مجموعة تتعلم المقاطع ذات التداعي العالي أولاً والأخرى تتعلم المقاطع ذات التداعي المنخفض أولاً. فذلك سوف "يوازن بين" أثار تعلم التعلم أو التعب. وعلى أية حال، إذا ساعد تعلم المقاطع ذات التداعي العالي، الأفراد على تعلم المقاطع ذات التداعي المنخفض لاحقاً، بينما العكس لا يكون صحيحاً، فإن ذلك يربك تفسير النتائج.

ضبط الفروقات الموقفية Controlling Situational Differences

إضافة إلى الفروقات بين الأفراد، من الضروري، أيضاً، ضبط أية متغيرات دخيلة قد تؤثر في الموقف التجريبي. فإذا كانت المتغيرات الموقفية غير منضبطة في تجربة معينة، فليس بوسع المرء أن يتأكد ما إذا كان المتغير المستقل، أو تلك الفروقات العرضية الفاعلة في المجموعات، هي التي تنتج الفرق في المتغير التابع. وهذه المتغيرات الدخيلة غير المنضبطة التي قد تعزى آثارها خطأً إلى المتغير المستقل تدعى "متغيرات الإرباك أو الإفساد / التشويش".

مثلاً، دعنا نفترض أن باحثاً يهتم بتأثير فيلم في إحداث تغييرات في الاتجاه نحو قضية معينة. يتم اختيار مجموعة واحدة من الأطفال عشوائياً من صف الانطلاق كي ترى الفيلم، بينما تترك مجموعة المقارنة في الصف. وقد يعرض الباحث، دون قصد، عدداً كبيراً من مواقف العنف. وقد يستاء الأطفال في المجموعة الضابطة، أو يشعرون بالرفض، أو الدونية، بالنسبة للآخرين. فقد يكون لأي من هذه العوامل أثر على نتيجة الدراسة. إن الفرق الذي يريد الباحث أن يعزوه إلى استخدام الفيلم، قد يعزى حقاً لأحد هذه الجوانب العرضية. وفي هذه الحالة يجب اتخاذ خطوات لتأمين قيام أفراد المجموعة الضابطة برؤية فيلم من نوع ما، أو لا تعلم أي من المجموعتين أنها تساهم في تجربة. إن هذا الحذر ضروري وذلك لضبط تأثير ظاهرة هوثورن.

من المهم جداً ضبط اتجاهات الأفراد في التجارب على الأدوية. مثلاً، في دراسة تأثير الدواء على أداء مهارة يدوية يجب أن يظن جميع الأفراد بأنهم يتناولون الدواء. ويتم ذلك من خلال استخدام عقار وهمي (Placebo)، الذي هو مادة خاملة / غير فعالة أو حيادية، يعطى إلى الأفراد بغية جعلهم يعتقدون أنهم يحصلون على الدواء التجريبي. ويعطي الباحث كل فرد مادة تبدو متماثلة تماماً. فبالنسبة لبعض الأفراد يعتبر ذلك هو الدواء قيد البحث، بينما للبقية هو الدواء الوهمي. وهكذا فإن توقع أثر الدواء يعد ثابتاً بين أفراد المجموعة التجريبية والضابطة، وعليه يستبعد أن يكون متغيراً مربكاً / مشوّشاً. وخلافاً لذلك، فإن مجرد المعرفة بأنهم يتناولون دواءاً، قد يجعل أفراد المجموعة التجريبية أكثر حذراً أو طيشاً - فلا يعرف الباحث ما إذا كان أثر الدواء أو سلوك الأفراد أو كلاهما هو الذي أدى إلى النتيجة.

إن استخدام الدواء الوهمي كما وصف أعلاه يوضح ما يدعى بالتجربة "أحادية التعمية" (Single - blind)، فالأفراد لا يدركون ظرف المعالجة التي هم فيها، مع أن الباحث يعرف. وأحياناً يكون من الضروري جعل اتجاهات الباحث ثابتة بالنسبة لمستويات المتغير المستقل المختلفة. ويتم ذلك باستخدام إجراء / أسلوب التجربة "ثنائية التعمية" التي لا يعرف فيها الباحث ولا الأفراد أي نوع من المعالجة يتعرضون لها. وتعد، التجارب "ثنائية التعمية" (double - blind)، أصعب في التنفيذ لأن طبيعة المعالجة غالباً ما تجعل من الواضح معرفة المجموعة التجريبية. وثمة اعتبار آخر هو أنه ينبغي على الباحث في موقف ثنائية التعمية، أن يعتمد على أناس آخرين لتكوين المجموعات وإجراء المعالجة وتسجيل النتائج.

طرق ضبط المتغيرات الموقفية: هناك ثلاثة طرق شائعة الاستخدام لضبط المتغيرات الموقفية المشوشة: تثبيتها، أو توزيعها عشوائياً أو استخدامها بصورة منتظمة ومنفصلة عن المتغير المستقل الرئيسي.

ويعني تثبيت المتغيرات الدخيلة أن جميع الأفراد في المجموعات المختلفة يعالجون بصورة متماثلة تماماً، باستثناء تعرضهم للمتغير المستقل. مثلاً، في تجربة القراءة، يكون من الضروري

ضبط حجم المجموعات لأنه من المعروف أن حجم المجموعة عامل مؤثر على تحصيل القراءة. ويجب أن يرى المرء بأن للمجموعات التجريبية والضابطة العدد نفسه من الأفراد. أما متغير المدرس، فيجب ضبطه أيضاً. لأن كفاءة وحماسة المدرس عاملان قد يؤثران على نتيجة أي تجربة تعلم. وهكذا ينبغي استخدام المدرس ذاته في طريقتي التدريس اللتين تتم مقارنتهما. ففي تجربة معينة، يجب أن يتبع المساعدون الإجراءات ذاتها: استخدم ذات التعليمات والأجهزة والاختبارات واحرص على إبداء الاتجاهات ذاتها مع كل المجموعات. وينبغي على جميع المجموعات أن تلتقي في الوقت ذاته من اليوم، وفي النوع ذاته من الغرف. وقد لا يود المرء أن تلتقي المجموعة التجريبية خلال الحصة الأولى في الصباح من يوم مدرسي، بينما تلتقي المجموعة الضابطة خلال الحصة الأخيرة من اليوم. إن الظروف البيئية - كدرجة الحرارة، وشدة الضوء، والرطوبة، وأثاث الغرفة، ووجود أو غياب الضوضاء، ينبغي لها أن تكون متماثلة لجميع المجموعات.

إذا لم يتسن تثبيت الظروف، فيجب على الباحث أن يحاول العشوائية أو موازنة بعض المتغيرات الموقفية. مثلاً، إذا تعذر الحصول على المدرس ذاته لكلا المجموعتين، فقد يخصص الباحث بصورة عشوائية نصف المدرسين المتيسرين لاستخدام الطريقة (أ) والنصف الآخر للطريقة (ب)، ويمكن اتباع الشيء ذاته مع الظروف التجريبية الأخرى كالأجهزة. وبهذه الطريقة يتم توزيع المتغيرات الموقفية عشوائياً، فأنواع من الظروف الدخيلة يتم تمثيلها دون أن يسمح لها بالتأثير على المتغير التابع بشكل منتظم.

وثمة طريقة أخرى لضبط المتغيرات الموقفية الدخيلة، وذلك بتفعيلها بشكل نظامي. ففي العديد من التجارب التربوية يكون من الضروري استخدام تسلسل لظروف تجريبية وضابطة، بغية ضبط الآثار المتنامية كتلك التي للممارسة والإرهاق. ويتم ذلك، بضبط الترتيب الذي تعرض فيه الظروف التجريبية من خلال الموازنة المضادة. فقد يتبع نصف الأفراد الترتيب AB والنصف الآخر الترتيب BA. وفي هذه الحالة يتم تفعيل المتغير الدخيل بشكل منتظم. ولا يضبط هذا الإجراء الأثر المشوش للترتيب فحسب، بل يمكن أن يوفر تقديراً لحجم أثر الترتيب بتحديد ما إذا كان متوسط قيم A، B في كلا الترتيبين مختلفاً.

وإذا اعتقد الباحث أن حجم الصف قد يؤثر على فاعلية طريقة التدريس الجديدة، فبوسعه أن يضبط ذلك المتغير بإدماجه في التصميم كمتغير مستقل ثان. وقد يقدم التصميم مستويين لمتغير الطريقة، ومستويين لحجم الصف كما هو مبين فيما يلي:

		حجم الصف	
		صغير	كبير
الطريقة	أ		
	ب		

ويمكن للمرء عندئذ أن يحدد أثر طريقة التدريس ذاتها، وأثر حجم الصف، وأي تأثير تفاعل بين الطريقة وحجم الصف. إن هذه الطريقة لضبط المتغيرات الدخيلة تماثل ذات الشيء، كإضافة متغيرات مستقلة أكثر إلى التجربة. ومع أنها تزيد تعقيد الدراسة، فإن لها مزية تقديم معلومات إضافية حول أثر المتغيرات ذات الصلة بالمتغير التابع وتفاعلها مع المتغيرات المستقلة. إن استخدام هذه الطريقة للضبط آخذ في الازدياد منذ دخول الحاسوب لمعالجة تحليل البيانات في الدراسات المعقدة. فتصميمات تحليل التباين ذي الطريقين وتصميمات تحليل التباين الأكثر تعقيداً تسمح باستقصاء متزامن لعدد من المتغيرات التي تدرس على انفراد وفي تفاعل - ويعد الأخير، الجانب الأهم للدراسة.

الصدق الخارجي لتصميمات البحوث

External Validity of Research Designs

المعيار الثاني للتصميم البحثي هو أن يكون له "صدق خارجي" الذي يشير إلى تعميم أو تمثيل الاستنتاجات. ويسأل الباحث سؤالاً معيناً وهو: لأية مجتمعات إحصائية، وأوضاع، ومتغيرات تجريبية، ومتغيرات قياس يمكن تعميم هذه الاستنتاجات؟

تتم أية دراسة، بالضرورة، على مجموعة من الأفراد مع أدوات قياس مختارة، وتحت ظروف فريدة في بعض الأوجه. إلا إن الباحث يريد أن توفر نتائج دراسته معلومات حول حقل واسع من الموضوعات / الأفكار، والظروف، والعمليات، أوسع مما قد تم البحث فيه أصلاً. ولإجراء التعميمات من المرصود إلى غير المرصود، نحتاج إلى تقدير مدى تمثيل عينة الأحداث المدروسة فعلاً لمجتمع واسع يمكن تعميم النتائج عليه. وحسب المدى الذي يمكن تعميم نتائج التجربة فيه على مختلف الأفراد والأوضاع وأدوات القياس، فإن التجربة تمتلك صدقاً خارجياً.

لقد حدد (Smith & Glass, 1987) ثلاثة أنماط من الصدق الخارجي: الصدق الخارجي للمجتمع الإحصائي، الصدق الخارجي البيئي، والصدق الخارجي للعمليات. إن الصدق الخارجي للمجتمع الإحصائي يتعلق بتحديد المجتمع الذي تعمم عليه نتائج التجربة. وهذا يطرح السؤال التالي: ما هو مجتمع الأفراد الذين يتوقع لهم أن يتصرفوا بذات الطريقة التي تصرف بها أفراد العينة؟ بينما يتعلق الصدق الخارجي البيئي بتعميم الآثار التجريبية في ظروف بيئية أخرى. وهذا يطرح السؤال التالي: في أية أوضاع يمكن توقع النتائج ذاتها؟ أما الصدق الخارجي للعمليات فيتعلق بمدى تمثيل التعريفات الإجرائية والإجراءات التجريبية، للمفاهيم البنائية ذات الاهتمام، فيكون السؤال: هل سيجد باحث آخر، العلاقات ذاتها، إن قام ببحث المسألة ذاتها، مستخدماً عمليات مختلفة.

الصدق الخارجي للمجتمع الإحصائي Population External Validity

إن أمل الباحث، هو إمكانية تعميم الاستنتاجات، من أفراد المجموعة التجريبية، على مجتمع أكبر

بكثير لم تتم دراسته بعد. فمثلاً، دعنا نفترض أن باحثاً كان قد اكتشف طريقة جديدة فاعلة لتدريس القراءة إلى عينة من طلبة الصف الأول، ويود أن يستنتج إن هذه الطريقة أفضل للمجموعات الأخرى من طلبة الصف الأول، وربما لجميع طلبة الصف الأول في الولايات المتحدة. ولكي يكون الباحث قادراً على تعميم استنتاجات صادقة، من النتائج التجريبية، على مجتمعات أكبر، فإنه يجب أن يحدد بشكل صحيح المجتمعات التي يمكن تعميم النتائج عليها. وما هو وثيق الصلة بهذه المسألة هو تمييز (Kempthorne, 1961) بين مجتمع سهل المنال، تجريبياً" والمجتمع المستهدف، كما أشير إليهما في الفصل 6. ويشير مجتمع سهل المنال، تجريبياً، إلى مجتمع الأفراد الذين يمكن الوصول إليهم، أو المتوفرين للباحث في دراسته. والمجتمع المستهدف هو المجموعة الشاملة للأفراد التي يريد الباحث أن يطبق عليها نتائج الدراسة. وفي المثال السابق يحتمل أن يكون المجتمع التجريبي سهل المنال لجميع طلبة الصف الأول في مقاطعة المدارس المحلية. ويكون المجتمع المستهدف جميع طلبة الصف الأول في الولايات المتحدة.

تحدث تعميمات الباحثين في مرحلتين: 1) من العينة إلى المجتمع القابل للوصول إليه (سهل المنال)، تجريبياً و 2) من المجتمع سهل المنال إلى المجتمع المستهدف. فإذا اختار الباحث عشوائياً عينة من مجتمع قابل للوصول (طلبة الصف الأول في مقاطعة المدارس) فإن الاستنتاجات يمكن تعميمها على هذه المجموعة الكبيرة دون صعوبة. ويشير الإحصاء الاستدلالي إلى احتمال أن ما هو صحيح لعينة هو صحيح أيضاً للمجتمع الذي استمدت منه. وفي المرحلة الثانية يريد الباحث التعميم من المجتمع سهل المنال، أي الذي يمكن الوصول إليه، إلى المجتمع المستهدف (جميع طلبة الصف الأول في الولايات المتحدة). ولا يمكن إجراء هذا النوع من التعميم بدرجة الثقة ذاتها كالنوع السابق. وإجراء مثل هذا الاستنتاج، يتطلب معرفة عميقة بخواص كلا المجتمعين. فكلما كان التشابه أكبر بين المجتمعين، سهل المنال، والمستهدف كانت الثقة أكبر في التعميم من مجتمع إلى آخر. فالتعميم من المجتمع سهل منال إلى المجتمع المستهدف، لا يمكن إجراؤه إحصائياً. إنها مسألة وصول إلى حكم. ويجب أن يوصف المجتمع سهل المنال بشكل عميق قدر الإمكان، والإشارة إلى أية ناحية يختلف بها هذا المجتمع عن المجتمع المستهدف. فعندئذ يكون بوسع الباحث وقارئي البحث إصدار أحكام ذكية حول مدى احتمال إمكانية تعميم النتائج على المجتمع المستهدف أو أي مجتمع آخر.

وثمة جانب آخر للصدق الخارجي للمجتمع الإحصائي، هو إمكانية تفاعل خصائص الأفراد والمعالجة. فحينما لا يكون مجتمعان سهلا المنال ممثلين لذات المجتمع المستهدف فإن دراسات تبدو متشابهة يمكن أن تؤدي إلى نتائج مختلفة تماماً – فالعلاقة بين المعالجة والمتغير التابع قد تحدث مع مجموعة واحدة، ولا تحدث مع مجموعة أخرى ذات خصائص مختلفة. لذا، لن يكون بالإمكان تعميم الاستنتاجات من مجموعة إلى أخرى. فطريقة الإرشاد (أ) قد تعطي نتائج أفضل من الطريقة (ب) في المدارس داخل المدن، بينما الطريقة (ب) أفضل للمدارس على

أطراف المدن. إن الطريقة الأفضل لتدريس الاستنتاج الكمي لطلبة الصف الثاني قد تكون الطريقة الأسوأ لطلبة الصف الثامن. وكما يقول المثل القديم "طعام قوم، هو سم قوم آخرين". ومرة أخرى، فإن الوصف المتقن للمجتمع سهل المنال سيساعد التربويين الآخرين في الحكم ما إذا كانت معالجة معينة ستعتبر "طعاماً أم سماً" لمجتمعاتهم.

إن بوسع الباحثين في الغالب تعزيز الصدق الخارجي للدراسة بتقسيم الأفراد حسب خواص ذات صلة واستخدام التصميم العاملي (أنظر الفصل 6). فمثلاً، مع مجتمع مدرسة متوسطة، سهل المنال، لدراسة تقارن فاعلية طريقتين في التدريس يمكن للمرء إجراء تصميم عاملي هو 2 × 3 × 2 (طرق × مستويات سنوات الدراسة × الجنس).

ولن تمكن النتائج الباحث من مقارنة فاعلية الطريقتين فقط، بل كذلك رؤية مدى ثبات النتائج بين طلبة الصف السادس، والسابع، والثامن، وكذلك بين البنين والبنات.

الصدق الخارجي البيئي Ecological External Validity

يجب أن يهتم المجربون أيضاً بالصدق الخارجي البيئي، بمعنى أنهم يحتاجون أن يكونوا قادرين على القول بأن المرء يتوقع الحصول على ذات الاستنتاجات في أوضاع أخرى. وللحصول على الصدق البيئي يجب أن يوفر التصميم التأكيد على أن الأثر التجريبي مستقل عن البيئة التجريبية المعنية.

من الواضح إن المطلب الأول للصدق للبيئي هو قيام الباحث بتقديم وصف تام للبيئة التجريبية المعنية في الدراسة. وعندئذ فقط، يستطيع القارئ الحكم على المدى الذي يمكن فيه للنتائج أن تعمم في أوضاع أخرى. مثلاً، هل تعمم النتائج الحاصلة من صف واسع مجهز جيداً على صف مزدحم سيئ التجهيز؟ هنا مرة أخرى، ينبغي على الباحث أن يقدم وصفاً كاملاً للبيئة التي تمكن القارئ في إعطاء الحكم المعقول الخاص بتعميم النتائج. وغالباً ما تجرى التجارب في بيئات يمكن فيها ضبط متغيرات مثل الضوء والضوضاء والمشوشات الأخرى. ومثل هذا الضبط مطلوب لتعزيز الصدق الداخلي. وعلى أية حال، فإن المدى الذي يكون فيه مثل هذا الضبط للبيئة غير ممثل للبيئة الطبيعية للأفراد، يقلل من الصدق الخارجي للتجربة. وقد يكون مرغوباً القيام أولاً بمعالجة في بيئة لها ضبط كاف لتوفير صدق داخلي عال، وإذا تم ذلك بنجاح، فحاول ذلك مرة أخرى في بيئة طبيعية أكثر.

قد يكون هناك "أثر رد فعل" بسبب الترتيبات التجريبية، فمعرفة الأفراد أنهم يشاركون في تجربة، قد يؤثر على استجاباتهم للمعالجة. وأحياناً، فإن وجود مراقبين أو أجهزة خلال التجربة، قد يغير من الاستجابات الاعتيادية للأفراد المساهمين في التجربة بحيث لا يستطيع المرء التعميم المتعلق بأثر المتغير التجريبي، على أشخاص تعرضوا له في بيئة غير تجريبية.

إن تأثير / ظاهرة هوثورن الموصوف سابقاً هو مثال واحد عن تأثير رد الفعل باعتباره

مشكلة صدق خارجي بيئي. فخطر تأثير هوثورن يمكن ضبطه بالترتيب للمجموعة الضابطة بتجريب نوع معين من التدخل. وعلى أية حال، ينبغي أن لا يكون هذا التدخل نشاطاً له صلة أو ارتباط بالمتغير التابع في الدراسة. فمثلاً، في الوقت الذي تخضع فيه المجموعة التجريبية لتجربة المتغير المستقل الحقيقي للدراسة، تكون المجموعة الضابطة مشغولة في نشاط مثير آخر. ولأن كلا من المجموعة التجريبية والضابطة ستتأثر بظاهرة هوثورن فإن الظروف ستكون أكثر قابلية للمقارنة ويستطيع الباحث عزل التأثير الحقيقي للمتغير المستقل عن ظاهرة هوثورن.

وثمة مشكلات أخرى يمكن لها أن تؤثر على الصدق البيئي. فمثلاً، طريقة تدريس جديدة قد تبدو أكثر نجاحاً من القديمة، بينما، في الواقع، كلاهما مؤثر بصورة متماثلة لأن المدرسين والطلبة في البرنامج الجديد قد يشعرون أنهم استثنائيون ويبدون ويطورون حماسة لما يقومون به، بينما يشعر أولئك الذين في المجموعة الضابطة بأنهم مهملون. ومن ناحية أخرى، قد تبدو الطريقة الجديدة أقل فاعلية من القديمة، لأن المدرسين والطلبة يستخدمون مواد وإجراءات غير مألوفة ولم يتعلموا، بعد، استخدامها بصورة فاعلة في الوقت الذي يجرى فيه قياس المتغير التابع. مثلاً، عندما لخص (Schoen, 1976) نتائج بحث سابق يقارن نتائج التدريس الفردي مقابل التدريس الجماعي في مادة الرياضيات بمدرسة ابتدائية، وجد إن هناك سبع عشرة دراسة ذات نتائج دالة إحصائياً تفضل التدريس الجماعي وإحدى عشرة أبقت على الفرضية الصفرية، وفقط ثمان ذات فروقات دالة إحصائياً تفضل طريقة التعليم الفردي. إن أغلب الباحثين قد افترضوا أن التدريس الفردي سيكون أكثر تأثيراً. وقد يكون ذلك دليلاً على أن التدريس الجماعي لمعظم المجتمعات يكون أكثر فاعلية. وعلى أية حال، من الممكن أن يكون أداء الطلبة في التدريس الجماعي أفضل، لأنهم ومدرسيهم عرفوا كيفية مواكبة التدريس الجماعي، بينما مع التدريس الفردي، كان عليهم أن يتعلموا طريقة جديدة للإجراء إضافة إلى محتوى وحدة الحساب.

وقد تهدد بعض "تأثيرات التفاعل" تعميم الاستنتاجات التجريبية. مثلاً، قد يزيد الاختبار القبلي أو يقلل من حساسية أفراد التجربة أو استجاباتهم لمتغير التجريبي مما يجعل النتائج الناجمة عن مجتمع الاختبار القبلي غير ممثلة لآثار المتغير التجريبي بالنسبة لمجتمع لم يخضع لامتحان قبلي حيث تم اختيار أفراد التجربة منه. في هذه الحالة، بوسع المرء التعميم على المجموعات التي أجرت الاختبار القبلي وليس التي لم تخضع له.

الصدق الخارجي للعمليات External Validity of Operations

يجري الباحثون دراسات باستخدام تعريفات إجرائية محددة للمتغيرات المستقلة والتابعة وكذلك للإجراءات المحددة. إن السؤال الخاص بالصدق الخارجي للعمليات يسأل ما إذا كانت نتائج مماثلة يمكن توقعها من باحثين مختلفين ومع تعريفات إجرائية مختلفة و/أو إجراءات مختلفة.

إن أحد الجوانب التي تؤخذ بنظر الاعتبار، عندئذ، هو التعريفات الإجرائية للمتغيرات

المستقلة والتابعة. ففي بعض التجارب، يستثار القلق بواسطة الصدمة الكهربائية، وفي تجارب أخرى بالتعليمات الشفوية للأفراد. ولكن، ما المدى الذي يمكن فيه تعميم، استنتاجات دراسة تستخدم تعريفاً ما، على دراسة أخرى؟ هل الإحباط الناتج عن منع الأطفال من لعب مرغوبة هو ذات الإحباط الناتج عن إعطائهم مسائل لا يمكن حلها؟ وإذا استخدم اختبار موضوعي لقياس المتغير التابع، فهل بوسع الباحث القول بأن التأثير نفسه سيلاحظ في حالة استخدام اختبار المقالة كأداة قياس؟ وما مدى تمثيل درجات، اختبار تورنس للتفكير الإبداعي، للمفهوم البنائي للإبداع كما يعرّفه مدرس الإنجليزية أو الفنون؟ ويقيّم القارئ لتقرير بحث، هذا الجانب بالسؤال أولاً، عن مدى وضوح تعريف المتغيرات المستقلة والتابعة. فلأجل جميع الأهداف العملية، تعد التقارير ذات التعريفات الغامضة عديمة القيمة. وثانياً، يجب على القارئ أن يقيم ثبات وصدق هذه التعريفات وإجراء حكم يتعلق بمدى ملاءمة التعريفات الإجرائية لتعريفات المفاهيم البنائية التي لها معنى لديه.

وثمة مشكلة صدق خارجي محتمل آخر وهي "حساسية الاختبار القبلي ". فإجراء الاختبار القبلي قد يغير أو يثير حساسية الأفراد، للاستجابة إلى المعالجة، بطريقة مختلفة عما هي عليه فيما إذا لم يجري اختبارهم قبلاً. مثلاً يعطى، لمجموعة من طلبة الصف السابع، استبيان يخص عادات الأكل لديهم، ويتم تقسيمهم إلى مجموعة تجريبية ومجموعة ضابطة. وتشاهد المجموعة التجريبية سلسلة من عروض الأفلام التي تخص عادات الأكل الجيدة، بينما تشاهد المجموعة الضابطة سلسلة من أفلام الصحة التي ليس لها صلة بعادات الأكل (عروض زائفة). ويستمد المتغير التابع عن طريق رصد اختيارات الأطفال للطعام في موقف فعلي ذي اختيار حر. فإذا أبدت المجموعة التجريبية تفضيلاً أكبر للأغذية الصحية، فإن الباحث سيميل إلى الاستنتاج بأن الأفلام فاعلة. وقبل التوصل إلى استنتاج، يجب على الباحث أن يدرس احتمال أن يكون الاختبار القبلي قد جعل الطلبة يفكرون بشأن عادات الأكل لديهم، و "هيأهم" للاستجابة إلى الأفلام. وقد لا يلاحظ التأثير ذاته في مجموعة لم تأخذ الاختبار القبلي. ولاحقاً، سنبحث طرق عزل أو تجنب تفاعل معالجة الاختبار القبلي (ويشمل عدم إجراء اختبار قبلي، والاعتماد على توزيع عشوائي للحصول على مجموعات تكون متماثلة تقريباً في الاختيار القبلي إذا ما طبق). لاحظ أن حساسية الاختبار القبلي تختلف عن إجراء الاختبار كتهديد للصدق الداخلي. ويشير الأخير إلى الاختبار بحد ذاته مما يسبب فرقاً في المتغير التابع. أما الأول فإنه يعني بأن الاختبارات القبلية والمعالجات معا تنتج تأثيراً قد لا تستطيع المعالجة وحدها إنتاجه.

ورغم أن الصدق الداخلي أمر لابد منه، فإن الباحث يحتاج إلى اختيار تصميم متين في كل من الصدق الداخلي والخارجي. وعلى أية حال، ففي بعض الحالات يميل الحصول على نوع معين من الصدق إلى تهديد الأنواع الأخرى. فمثلاً، عندما نرتب ضبطاً صارماً أشد، في تجربة تربوية، قد نزيد من زيفها ونقطع سبيل قابلية تطبيق الاستنتاجات على وضع صفي فعلي.

ومن ناحية عملية، نحاول الوصول إلى حل وسط بين الصدقين الداخلي والخارجي، بما يرقي إلى اختيار تصميم يوفر ضبطاً كافياً لجعل النتائج قابلة للتفسير مع الاحتفاظ ببعض الواقعية لكي يتم تعميم الاستنتاجات على الأوضاع المطلوبة.

تصنيف التصميمات التجريبية
CLASSIFICATION OF EXPERIMENTAL DESIGNS

تصنف التصميمات، حسب مناقشة التصميمات التالية، إلى: "قبل التجريبية" أو "تجريبية حقيقية" أو "شبه تجريبية" اعتماداً على درجة الضبط المتوفرة. فليس للتصميمات قبل التجريبية تخصيص عشوائي للأفراد في مجموعات أو أي استراتيجيات أخرى لضبط المتغيرات الدخيلة. أما التصميمات التجريبية الحقيقية، فتستخدم العشوائية وتقدم أقصى ضبط للمتغيرات الدخيلة. وتفتقر التصميمات شبه التجريبية إلى العشوائية، لكنها تستخدم استراتيجيات أخرى لتأمين بعض الضبط للمتغيرات الدخيلة. وهي تستخدم مثلاً، عندما تكون الصفوف كمجموعات تجريبية وضابطة. وهكذا، فالتصميمات التجريبية الحقيقية لها أكبر صدق داخلي بينما لشبه التجريبية صدق أقل وللتصميمات قبل التجريبية أقل صدق داخلي.

من الضروري، قبل البدء بمناقشة التصميمات التجريبية، تقديم المصطلحات والرموز التي سيتم استخدامها:

1- X، يمثل المتغير المستقل الذي يفعّله الباحث، وسوف يشار إليه بالمتغير التجريبي أو المعالجة.

2- Y، يمثل مقياس المتغير التابع. Y_1، يمثل المتغير التابع قبل تفعيل المتغير المستقل X، وهو عادة ما يكون اختباراً قبلياً لنوع معين يجرى قبل المعالجة التجريبية. Y_2 يمثل المتغير التابع بعد استخدام / تفعيل المتغير المستقل X، وهو عادة ما يكون اختباراً بعدياً يجرى على الأفراد بعد المعالجة التجريبية.

3- S، يمثل الفرد أو المستجيب المستخدم في التجربة، والجمع هو Ss (كل الأفراد).

4- E، يشير إلى المجموعة التجريبية – أي المجموعة التي تعطى معالجة المتغير المستقل.

5- C، يشير إلى المجموعة الضابطة – أي المجموعة التي لم تحظ بالمعالجة التجريبية. أنها تحظى بمعالجة مختلفة، أولاً تحظى بأية معالجة على الإطلاق.

6- R، يوضح التخصيص العشوائي للأفراد في المجموعات التجريبية والتخصيص العشوائي للمعالجات في هذه المجموعات.

7- M، يوضح إن الأفراد قد تمت مطابقتهم زوجياً ثم يخصص أفراد كل زوج إلى مجموعات المقارنة، عشوائياً.

وفي نماذج الأمثلة (Paradigms)، للتصميمات المختلفة، فإن المتغيرات المستقلة (Xs) والتابعة (Ys) الموجودة في صف معين r تطبق على الأفراد أنفسهم. ويوضح البعد، من اليسار إلى اليمين، الترتيب الزمني، بينما Xs و Ys المتعامدة مع بعضها فتقدم بشكل متزامن. وتوضح الفاصلة (–) أن المجموعة الضابطة لم تحظ بالمعالجة X.

التصميمات قبل التجريبية Preexperimental Designs

يقدم هذا الجزء تصميمين، ثم تصنيفهما على أنهما قبل تجريبيين، لأنهما يوفران القليل أولا يوفران أي ضبط للمتغيرات الدخيلة. ولسوء الحظ، يجد المرء أن هذه التصميمات لا تزال قيد الاستخدام أحياناً في البحوث التربوية. وسيكون من المفيد بدء نقاشنا مع هذه التصميمات الضعيفة لأنها توضح جيداً الطريقة التي تعمل فيها المتغيرات الدخيلة التي تعمل على تعريض الصدق الداخلي للتصميم إلى الخطر. وإذا أصبح القارئ مدركاً لمصادر الضعف هذه في التصميم، فلا بد له أن يكون قادراً على تجنبها.

التصميم 1: تصميم المجموعة الواحدة ذو الاختبارين القبلي والبعدي

Design 1: One-Group Pretest-Posttest Design

يشمل تصميم المجموعة الواحدة، عادة، ثلاث خطوات: 1) إجراء اختبار قبلي لقياس المتغير التابع، 2) تطبيق المعالجة التجريبية X على الأفراد، 3) إجراء اختبار بعدي وقياس المتغير التابع مرة أخرى. ثم يجري تقييم الفروقات التي تعزى إلى تطبيق المعالجة التجريبية عن طريق مقارنة درجات الاختبارين القبلي والبعدي

التصميم 1: تصميم المجموعة الواحدة للاختبارين القبلي والبعدي

الاختبار البعدي	المتغير المستقل	الاختبار القبلي
Y_2	X	Y_1

لتوضيح استخدام هذا التصميم، افترض أن مدرس ابتدائية يود تقييم فاعلية أسلوب جديد لتدريس طلبة الصف الرابع، دراسات اجتماعية. ففي بداية السنة الدراسية يعطي الطلبة اختباراً مقنناً يبدو مقياساً جيدا لتحصيل أهداف الدراسات الاجتماعية للصف الرابع. ثم يدخل المعلم أسلوب التدريس الجديد، وفي نهاية السنة يجري الاختبار المقنن مرة أخرى، ويقارن درجات الاختبارين الأول والثاني لغرض تحديد فرق التعرض لطريقة التدريس الجديدة (X).

ولما كان التصميم يشمل مجموعة واحدة ومدرساً واحداً، فإنه يبدو ضابطاً للفروقات بين الأفراد، ومتغيرات الموقف. أما الضبط فهو مصطنع على أية حال.

إن القصور الكبير لتصميم المجموعة الواحدة، هو أنه لا يمكن للباحث، بسبب عدم استخدام مجموعة ضابطة، أن يفترض أن التغير بين الاختبارين القبلي والبعدي، أحدثته المعالجة التجريبية. فهناك دائماً احتمال قيام بعض المتغيرات الدخيلة، تسببت بالتغير كله أو بجزء منه. وهكذا، يفتقر هذا التصميم إلى الصدق الداخلي.

ما هي بعض المتغيرات الدخيلة التي يمكن أن تعمل على إحداث التغيير الملاحظ بين درجات الاختبار القبلي والاختبار البعدي؟ هناك متغيران دخيلان واضحان لا يمكن ضبطهما في هذا التصميم، وهما "التاريخ والنضوج". ويشير التاريخ، باعتباره مصدراً للتباين الدخيل، إلى أحداث خاصة يمكن لها أن تحدث بين الاختبار القبلي والاختبار البعدي، وليس للمعالجة التجريبية. وفي مثال الدراسات الاجتماعية، فإن اهتمام المجتمع المحلي الواسع بالانتخابات، زاد التأكيد على الدراسات الاجتماعية في المدرسة، أو استقدام مدرس مؤثر، على وجه الخصوص، يمكن أن يزيد من تحصيل الطالب في هذا المجال. وقد تتسبب عدوى، بغياب متزايد عما يقلل من التحصيل. أما النضوج فيشير إلى التغيرات، في الأفراد أنفسهم، التي تحدث مع مرور الوقت. وفيما بين الاختبار القبلي والاختبار البعدي ينمو الأطفال ذهنياً وبدنياً، وقد يتعلمون تجارب يمكن لها أن تؤثر على المتغير التابع. ويصبح التاريخ والنضوج مصدرين مؤثرين بشكل متزايد على التباين الدخيل عندما تكون الفترة الزمنية بين Y_1 و Y_2 طويلة. أما اتجاهات الأفراد، والتنفيذ، والارتداد فهي أخطار، لا يمكن ضبطها، للصدق الداخلي.

عيب آخر للتصميم 1، هو أنه لا يقدم، أية طريقة لتقييم تأثير الاختبار القبلي نفسه. ونعلم إن هناك، تأثير الممارسة، عندما يأخذ الأفراد الاختبار للمرة الثانية أو حتى عندما يأخذون شكلاً بديلاً للاختبار - فالأفراد يؤدون بشكل أفضل في المرة الثانية، حتى بدون أية تعليمات أو نقاش خلال الفترة. وهذا صحيح ليس فقط لاختبارات التحصيل والذكاء، بل لاختبارات الشخصية أيضاً. وفي حالة الاختبارات الشخصية نلاحظ ميلاً لحو تكيف أفضل، على وجه العموم.

والكسب في إعادة الاختبار، هو جانب من جوانب المشكلة الكبرى لتفاعلية أدوات القياس. وتشير التفاعلية إلى حقيقة أن هناك في الغالب تفاعل بين الفرد ومقياس الاختبار القبلي وأن هذا التفاعل، وليس تفعيل X، هو الذي يولد التغيير في قياس Y_2. والمقاييس التي تسبب قيام الفرد برد فعل تدعى "المقاييس التفاعلية". مثلاً، في دراسة لتغيير الاتجاهات قد يعمل قياس الاتجاهات بذاته كحافز، أي أن الفرد قد يقوم برد فعل إزاء مضمون المقياس، وأن هذا التفاعل هو الذي يسبب التغير المرصود / الملاحظ في الاتجاهات، حتى بدون أية معالجة تجريبية. وهذا التأثير يكون واضحاً جداً حينما يكون مضمون/ محتوى الاختبار القبلي جديداً أو مثيراً للجدل، أو عندما يكون له تأثير محفز للأفراد بشكل خاص.

إن للتصميم 1، القليل مما يوصي به. فدون مجموعة ضابطة لإجراء المقارنة الممكنة، فإن النتائج المتوفرة من تصميم المجموعة الواحدة غير قابلة للتفسير أساساً.

التصميم 2: مقارنة المجموعة الثابتة Design 2: Static Group Comparison

يستخدم التصميم 2، اثنتين أو أكثر من مجموعات موجودة قبلاً أو تامة (ثابتة)، وواحدة منها فقط، هي التي تتعرض للمعالجة التجريبية. ولا يتم تخصيص الأفراد بشكل عشوائي للمجموعات. فالباحث يفترض، فقط، أن المجموعات متكافئة في جميع الجوانب ذات الصلة، وإنها قد تختلف في تعرضها إلى X، فحسب. وتقارن قياسات المتغير التابع للمجموعات لتحديد تأثير تعرضها للمعالجة X. ويستخدم، أحياناً، هذا التصميم في طرق البحث في التربية. ويقارن تحصيل الطلبة الذين تم تدريسهم حسب الطريقة الجديدة، مع تحصيل صف مماثل تم تدريسه حسب الطريقة القديمة.

إن للتصميم 2 مجموعة أو مجموعات ضابطة، تسمح بالمقارنة المطلوبة الجديرة بالاحترام العلمي. وعلى أية حال، هناك عيب أساسي في هذا التصميم. فحيث لا يتم استخدام العشوائية، ولا حتى المزاوجة لتخصيص الأفراد للمجموعات التجريبية والضابطة، لا يمكننا أن نفترض أن المجموعات متكافئة قبيل المعالجة التجريبية، وهي قد تختلف في بعض المتغيرات ذات الصلة، وقد تكون هذه الفروقات، وليس X، هي المسؤولة عن التغير الملاحظ. ولأننا لا نستطيع أن نتأكد أن المجموعات متماثلة بالنسبة لجميع العوامل التي قد تؤثر على المتغير التابع، فإن هذا التصميم يعتبر مفتقراً للضبط الضروري، ويجب تصنيفه على أنه قبل تجريبي.

التصميم 2: مقارنة المجموعة الثابتة

المجموعة	المتغير المستقل	الاختبار البعدي
E	X	Y_2
C	–	Y_2

التصميمات التجريبية الحقيقية True Experimental Designs

تدعى التصميمات في هذا الصنف بالتجارب الحقيقية، لأن الأفراد يخصصون بشكل عشوائي للمجموعات. وهي تصميمات يوصى بها بشكل عال للتجارب في مجال التربية وذلك بسبب الضبط الذي توفره.

التصميم 3: تصميم المجموعة الضابطة حسب تخصيص عشوائي للأفراد واختبار بعدي فقط
Design 3: Randomized Subjects, Posttest - Only Control Group Design

يعتبر التصميم 3 هو الأبسط، رغم كونه أقوى التصميمات التجريبية. فهو يتطلب مجموعتين من الأفراد يتم تخصيصهم عشوائياً، وتحدد لكل واحدة ظرف مختلف. ولا يستخدم أي اختبار قبلي.

ويضبط التوزيع العشوائي جميع المتغيرات الدخيلة المحتملة كما نؤكد بأن أية فروقات أولية بين المجموعتين، تعزى إلى الصدفة فقط، وعليه فإنهما تتبعان قوانين الاحتمال.

وبعد تخصيص الأفراد في المجموعتين، فإن المجموعة التجريبية تُعرّض إلى المعالجة التجريبية. وتعامل المجموعتان بصورة متماثلة من جميع الأوجه الأخرى. ثم يقاس أفراد كلا المجموعتين حسب المتغير التابع Y_2. وتقارن الدرجات لتحديد تأثير X. فإذا كان اختلاف الوسطين الناتجين للمجموعتين دالاً (أي أكثر اختلافاً مما هو متوقع على أساس الصدفة وحدها) فإن بوسع الباحث أن يكون على ثقة بأن الظروف التجريبية مسؤولة عن النتيجة المرصودة / الملاحظة.

والفائدة الرئيسية للتصميم 3 هي العشوائية، التي تؤكد على التكافؤ الإحصائي للمجموعتين قبيل إدخال المتغير المستقل. تذكر أنه كلما أزداد عدد الأفراد، أزداد احتمال العشوائية بتوفير مجموعتين متكافئتين. ونحن نوصي أن يكون 30 فرداً في كل مجموعة، على الأقل. ويقوم التصميم 3 بضبط الآثار الرئيسية للتاريخ، والنضوج، والارتداد، والاختبار القبلي، ونظراً لعدم استخدام أي اختبار قبلي، فلا يمكن أن يكون هناك تأثير تفاعلي للاختبار القبلي و X. ويوصى بهذا التصميم بشكل خاص، في الأوضاع التي يحتمل أن يحدث فيها تفاعل الاختبار القبلي. وهو مفيد في الدراسات التي لا يتوفر فيها الاختبار القبلي أو لا يكون مناسباً – كما هو الحال مثلاً في الدراسات الخاصة برياض الأطفال أو الصفوف الأولى، حيث يستحيل إجراء اختبار قبلي لأن التعلم لم يتضح بعد. وثمة فائدة أخرى لهذا التصميم، هي إمكانية توسيعه ليشمل أكثر من مجموعتين عند الضرورة. أما المشكلات المحتملة فهي اتجاهات الأفراد وأخطار التنفيذ التي قد تؤثر على الصدق الداخلي.

التصميم 3: تصميم المجموعة الضابطة حسب تخصيص عشوائي للأفراد واختبار بعدي فقط

	المجموعة	المتغير المستقل	الاختبار البعدي
(R)	E	X	Y_2
(R)	C	–	Y_2

ولا يسمح التصميم 3 للباحث أن يقيّم التغير. فإذا كان مثل هذا التقييم مطلوبا فإن تصميما (كالتصميم 5) الذي يستفيد من كل من الاختبار القبلي والاختبار البعدي سوف يتم استخدامه.

التصميم 4: تصميم المجموعة الضابطة حسب تناظر عشوائي للأفراد واختبار بعدي فقط

Design 4: Randomized Matched Subjects, Posttest-Only Group Design

يشابه هذا التصميم التصميم 3 عدا أنه يستخدم أسلوب التناظر وليس التخصيص العشوائي

للحصول على مجموعات متكافئة. ويتم تناظر الأفراد حسب متغير أو، أكثر من متغير، يمكن قياسه بشكل مناسب، مثل I Q أو درجة القراءة. وبالطبع فإن متغيرات التناظر المستخدمة هي تلك التي لها ارتباط دال(*) مع المتغير التابع. ورغم أن الاختبار القبلي غير مشمول بالتصميم 4، فإنه إذا تيسرت درجات اختبار قبلي حسب المتغير التابع فإن بالإمكان استخدامها بصورة فاعلة في إجراء التناظر. فالقياسات تجرى مزاوجتها بحيث أن درجات الأفراد المتناظرين تكون قريبة من بعضها قدر الإمكان، ويخصص فرد من كل زوج عشوائيا لمعالجة والآخر للمعالجة الثانية. ويمكن استخدام رمي العملة النقدية لتحصيل هذا التخصيص العشوائي.

التصميم 4: تصميم المجموعة الضابطة حسب تناظر عشوائي للأفراد واختبار بعدي فقط

الاختبار البعدي	المتغير المستقل	المجموعة	
Y_2	X	E	(M_r)
Y_2	X	C	

وتعد مزاوجة التناظر مفيدة جدا في الدراسات التي تستخدم عينات صغيرة وحين يكون التصميم 3 غير مناسب. ويعتمد التصميم 3 تماماً على التخصيص العشوائي للحصول على المجموعات المتكافئة. ومع العينات الصغيرة فإن تأثير الصدفة وحده قد يؤدي إلى وضع تكون فيه المجموعات العشوائية مختلفة أساسا عن بعضها. ولا يقدم التصميم 3 أية ضمانة بان المجموعات الصغيرة قابلة للمقارنة فعلاً قبل تطبيق المعالجات. إن تصميم مزاوجة التناظر للأفراد، على أية حال، يفيد في تقليص المدى الذي تعزى فيه الفروقات التجريبية إلى الفروقات الأولية بين المجموعتين، أي انه يضبط الفروقات الموجودة قبلا بين الأفراد حسب متغيرات مرتبطة جدا بالمتغير التابع الذي صممت التجربة للتأثير فيه. إن التخصيص العشوائي للأزواج المتناظرة في المجموعات يعد إضافة إلى متانة هذا التصميم.

إن التصميم 4 عرضة للصعوبات التي ذكرناها آنفا فيما يخص المزاوجة كوسيلة ضبط. فيجب أن يكون التناظر لكل الأفراد المحتملين تاماً كما يجب أن يكون تخصيص الأفراد لكل زوج للمجموعتين محدداً بصورة عشوائية. فإذا تم استثناء فرد واحد أو أكثر بسبب عدم وجود مناظر مناسب فإن ذلك سيعمل على تحيز العينة. وعند استخدام التصميم 4 فإن من المهم إيجاد مناظر لكل فرد، حتى إذا كان بصورة تقريبية، قبل التخصيص العشوائي. ويمكن استخدام التصميم 4 مع مجموعتين أو أكثر عن طريق إيجاد فئات متناظرة وتخصيص فرد عشوائيا من كل فئة لكل مجموعة في التجربة.

(*) هذا يعني أن $(r \geq 0.50)$ مما سبق – (المراجع).

التصميم 5: تصميم المجموعة الضابطة حسب تخصيص عشوائي للأفراد واختبارين قبلي وبعدي

Design 5: Randomized Subjects, Prettest-Posttest Control Group Design

في التصميم 5، يخصص الأفراد للمجموعتين التجريبية والضابطة عن طريق التخصيص العشوائي، وتعطيان اختبارا قبليا حسب المتغير التابع Y. وتقدم المعالجة فقط لأفراد المجموعة التجريبية لمدة محددة، تقاس بعدها المجموعتان حسب المتغير التابع. ويحسب فرق المتوسط(*) (أو متوسط الفرق) بين الاختبار القبلي والاختبار البعدي (Y_2-Y_1) لكل مجموعة ثم تقارن درجات فرق المتوسط بغية التأكيد ما إذا كانت المعالجة التجريبية قد أحدثت تغييرا أكبر من الموقف الضابط. إن الدلالة الإحصائية للفرق في معدل التغييرات (والتي يتم إيجادها عندما يجري متوسط معدل التغيير للمجموعة الضابطة من متوسط التغيير للمجموعة التجريبية). قد يتحدد عن طريق الاختبار التائي (t) أو الاختبار الفائي (F). وهناك إجراء آخر أكثر دقة ألا وهو إجراء تحليل التباين الاقتراني ANCOVA باعتبار درجات الاختبار البعدي كمتغير تابع ودرجات القبلي كمتغير تباين اقتراني (Covariate).

التصميم 5: تصميم المجموعة الضابطة حسب تخصيص عشوائي للأفراد واختبارين قبلي وبعدي

الاختبار البعدي	المتغير المستقل	الاختبار القبلي	المجموعة	
Y2	X	Y1	E	(R)
Y2	–	Y1	C	(R)

وحقيقة أن المجموعة الضابطة لا تحظى بمعالجة تجريبية لا يعني أن أفراد المجموعة الضابطة لم يحظوا بأية خبرة على الإطلاق. وفي البحث حول طرق التدريس، يتم تدريس المجموعة الضابطة عموماً عن طريق الإجراء التقليدي أو الاعتيادي. وفي بعض تجارب التعلم، من المألوفة إعطاء المجموعة الضابطة نوعا من النشاط الذي لا صلة له، بين الاختبارين القبلي والبعدي، بينما تتلقى المجموعة التجريبية تدريباً محدداً للمهمة. ففي التجربة حول آثار دواء معين، يعطي المرء دواء وهمياً للمجموعة الضابطة دون أن يجعل أفرادها على علم بأنهم يتلقون معالجة مختلفة عن المجموعة التجريبية.

(*) فرق المتوسط (average difference): مصطلح يشير هنا إلى الفرق بين متوسطي الدرجات (Y_2)، (Y_1) لكل مجموعة ويساوي: $\frac{\Sigma Y_2}{n} - \frac{\Sigma Y_1}{n} = \frac{\Sigma (Y_2 - Y_1)}{n}$ حيث n عدد أفراد المجموعة
وبالتالي فإن فرق المتوسط (أي الفرق بين المتوسطين) يساوي متوسط الفرق (والأوضح لغوياً، متوسط الفروق). (المراجع)

وتسمح القياسات القبلية والبعدية، في التصميم5، للباحث بدراسة التغير، وغالباً ما يشار إليه بالتصميم الكلاسيكي لتجارب التغير. وتكمن القوة الرئيسية لهذا التصميم في العشوائية الأولية التي تؤكد على التكافؤ الإحصائي بين المجموعات قبيل إجراء التجربة، وإلى حقيقة أن ضبط الباحث للاختبار القبلي يتيح له تدقيقا إضافيا لتكافؤ المجموعتين حسب الاختبار القبلي، Y_1. وهكذا، فإن التصميم 5 مع توزيعه العشوائي، يضبط أغلب المتغيرات الدخيلة التي تهدد الصدق الداخلي. فمثلاً، يتعرض جميع أفراد المجموعتين التأثيرات التاريخ، والنضوج، والاختبار القبلي، وعليه، فإن أي فرق بين المجموعتين في مقياس Y، لا يحتمل لها أن تعزى إلى هذه العوامل. ويضبط الاختيار الفارقي/ التمييزي للأفراد والارتداد الإحصائي من خلال إجراء التوزيع العشوائي. ويمكن استخدام التصميم 5 مع أكثر من مجموعتين.

وينصب الاهتمام الرئيسي في استخدام التصميم 5 على الصدق الخارجي. ومما يدعو للسخرية، هو أن المشكلة تنبع من استخدام الاختبار القبلي الذي هو جانب أساسي للتصميم. وكما ذكرنا سابقا فقد يكون هناك تفاعل بين الاختبار القبلي والأفراد الذين بوسعه تغييرهم أو زيادة حساسيتهم بطرق معينة. ورغم أن كلا من المجموعتين E وC تأخذ الاختبار القبلي وقد تتعرض لتأثير الحساسية فإن ذلك قد يجعل أفراد المجموعة التجريبية يستجيبون للمعالجة X بطريقة معينة فقط بسبب حساسيتهم المتزايدة. والسؤال المهم هو هل أن تأثير X على أفراد المجموعة التجريبية يظل كما هو بدون التعرض إلى الاختبار القبلي؟ وكانت هذه المشكلة واضحة بشكل خاص في دراسات تغير الاتجاهات. ولدى تطبيق مقياس الاتجاهات الأول كاختبار قبلي، في دراسة ما، فإن بوسعه أن يثير اهتمام أو حساسية الأفراد إزاء قضايا أو مواد مشمولة بالمقياس. ثم عندما تطبق المعالجة التجريبية (محاضرة، فيلم أو ما شابه)، فقد يستجيب الأفراد بالقدر نفسه للمعالجة X كما هو الحال تماماً في استجابتهم لمجموعة احساساتهم المثارة إزاء القضايا والمعالجة التجريبية.

دعنا ندرس مثلاً آخر. افترض أن أحد المعايير لنجاح طريقة تدريس جديدة في الدراسات الاجتماعية للمدارس الثانوية هو عدد الطلبة الذين قالوا أنهم يقرأون صحفا مثل "وول ستريت جورنال". ولم يجر خلال المقرر نفسه تأكيد خاص على هذا المصدر الخاص، لكنه تم توفير هذه الصحيفة مع صحف أخرى أقل شهرة، نوعا ما، للطلبة، فإذا استخدمت الدراسة تصميماً لاختبارين قبلي وبعدي، فقد يحتوي الاختبار القبلي على بند مثل: "هل تقرا صحيفة وول ستريت جورنال لأجل الأخبار اليومية؟" فهذا السؤال وحده، قد يكون كافيا لإثارة حساسية الطلبة في المجموعة التجريبية إزاء تلك الصحيفة، ولذا، حين تتوفر خلال المقرر، فإنهم على الأغلب، سيلتقطونها من بين الصحف الأخرى. ونتيجة لذلك، فإن المجموعة التجريبية قد تبدي استخداما لصحيفة وول ستريت جورنال في الاختبار البعدي أكثر مما تبديه المجموعة الضابطة – ليس بسبب محتوى المقرر وحده، بل بسبب التأثير المزدوج لمحتوى المقرر والاختبار القبلي.

فصف جديد تم تدريسه بذات الطريقة، لكنه لم يجر الاختبار القبلي، وعليه فإن حساسيته لم تثر، قد لا يبدي اهتماما لصحيفة وول ستريت أكثر من المجموعة الضابطة.

إن مثل هذا التأثير يمثل تفاعلا بين الاختبار القبلي والمعالجة التجريبية. ولان الاختبار القبلي قد يزيد (أو ينقص) حساسية أو استجابة الأفراد لتفعيل X، فإن النتائج التي تم الحصول عليها لعينة الاختبار القبلي قد تكون غير ممثلة لتأثيرات المتغير التجريبي بالنسبة لمجتمع، لم يجر له اختبار قبلي، واختيرت منه المجموعة التجريبية. وهكذا، تكون لدينا مشكلة في قابلية التعميم، فقد نكون قادرين على تعميم الاستنتاجات التجريبية على المجموعات التي أخذت الاختبار القبلي وليس على المجموعات التي لم تأخذه. وهذا التفاعل بين الاختبار القبلي والمعالجة تهديد للصدق الخارجي.

وعلى الرغم من هذا العيب، فإن التصميم 5 يستخدم بشكل واسع، لأن تأثير التفاعل ليس مشكلة خطيرة في أغلب البحوث التربوية. فالاختبارات القبلية المستخدمة هي غالباً، اختبارات تحصيل من نوع ما، وعليه، ليس لها تأثير متميز لإثارة حساسية الأفراد المعتادين على مثل هذا الاختبار. وعلى أية حال، إذا كانت إجراءات الاختبار جديدة أو محفزة في تأثيرها، فإنه يوصى بان يختار الباحث تصميما لا يشمل الاختبار القبلي. وبدلاً من ذلك، إن شك المرء بان تأثير الاختبار القبلي قد يكون فاعلاً فإن من الممكن إضافة مجموعة أو مجموعات جديدة للدراسة – مجموعة لم تأخذ الاختبار القبلي. واقترح (Solomon, 1949) تصميمين يتغلبان على ضعف التصميم 5 وذلك بإضافة مجموعة أو مجموعات لم تأخذ الاختبار القبلي، وتقدم هذه التصميمات في التالي.

التصميم 6: تصميم سولومون ذو المجموعات الثلاث

Design 6: Solomon Three-Group Design

يستخدم أول تصميمات سولومون ثلاث مجموعات مع تخصيص عشوائي لأفراد المجموعات. ويمكن أن نرى أن السطرين الأولين من هذا التصميم يتطابقان مع التصميم 5. وعلى أية حال، فإن لتصميم سولومون ميزة استخدام مجموعة ضابطة ثانية وهو بذلك يتغلب على الصعوبة الملازمة للتصميم 5 – أي، التأثير التفاعلي للاختبار القبلي والتفعيل التجريبي. وهذه المجموعة الضابطة الثانية المشار لها (C_2) لم تأخذ الاختبار القبلي، بل تعرضت إلى المعالجة X. ثم يتم استخدام قياساتها Y_2 لتقييم تأثير التفاعل.

التصميم 6: تصميم سولومون ذو المجموعات الثلاث

الاختبار البعدي	المتغير المستقل	الاختبار القبلي	المجموعة	
Y_2	X	Y_1	E	(R)
Y_2	–	Y_1	C_1	(R)
Y_2	X	–	C_2	(R)

ويتم تحصيل تقييم تأثير التفاعل من خلال مقارنة الدرجات Y_2 للمجموعات الثلاث. ويتم إدخال درجات الاختبار البعدي فقط في التحليل. ورغم أن للمجموعة التجريبية وسطاً حسب (Y_2) أعلى بشكل دال من وسط المجموعة الضابطة (C_1)، إلا أنه لا يسعنا أن نكون واثقين بأن هذا الفرق يعزى إلى (X). فربما حدث ذلك بسبب الحساسية المتزايدة للأفراد بعد الاختبار القبلي وتفاعل حساسيتهم مع (X). وعلى أية حال، إذا كان وسط (Y_2) للمجموعة الضابطة الثانية (C_2) أعلى بشكل دال من وسط المجموعة الضابطة الأولى، فعندئذ يمكننا أن نفترض أن المعالجة التجريبية وليس تأثير "تفاعل الاختبار القبلي -X" قد ولّد الفرق لأن المجموعة الضابطة الثانية لم تأخذ الاختبار القبلي. وهذه المجموعة رغم حصولها على معالجة (X) تعمل كمجموعة ضابطة لذا أشير لها بالرمز (C_2)(*).

التصميم 7: تصميم سولومون ذو المجموعات الأربع

Design 7: Solomon Four - Group Design

يقدم التصميم 7 ضبطا أكثر فعالية وذلك بتوسيع التصميم 6 ليشمل مجموعة ضابطة أخرى. ولا تحظى هذه المجموعة الرابعة باختبار قبلي أو معالجة. ومرة أخرى فإن المجموعة الثالثة رغم حصولها على المعالجة (X) فإنها تعمل كمجموعة ضابطة.

التصميم 7: تصميم سولومون ذو المجموعات الأربعة

	المجموعة	الاختبار القبلي	المتغير المستقل	الاختبار البعدي
(R)	E	Y1	X	Y_2
(R)	C_1	Y1	–	Y_2
(R)	C_2	–	X	Y_2
(R)	C_3	–	–	Y_2

وتكمن قوة التصميم 7 في كونه يتمتع بمزايا العديد من التصميمات إضافة إلى مساهمته الفريدة. فهو يقدم ضبطا جيدا لمخاطر الصدق الداخلي. فأول سطرين (التصميم 5) يضبطان العوامل الخارجية كالتاريخ والنضوج والسطر الثالث (التصميم 6) يضبط تأثير تفاعل "الاختبار القبلي -X". وعندما يضاف السطر الرابع في التصميم7 يكون لدينا ضبط للتأثيرات القائمة المحتملة التي قد تحدث بين Y_1 وY_2. أما السطران الأخيران فيمثلان التصميم 3، وبهذا نكون فعلاً قد دمجنا تصميم المجموعة الضابطة بتوزيع عشوائي للأفراد الذين اخذوا الاختبارين القبلي والبعدي مع تصميم المجموعة الضابطة لأفراد بتوزيع عشوائي واخذوا فقط الاختبار البعدي. وإضافة إلى الإفادة

(*) نشير إلى أن مجموعة مقارنة تكون "ضابطة" عندما يحجب تأثير أحد العناصر الفاعلة في التصميم عنها. (المراجع)

من قوة كل تصميم على حدة، فإن لدينا ملامح التكرار المطابق الذي توفره التجربتان. ويستغل هذا الدمج المعلومات المقدمة عن طريق إجراء الاختبارين القبلي والبعدي ويوضح في الوقت ذاته كيف يؤثر الظرف التجريبي على مجموعة الأفراد (Ss) التي لم تأخذ الاختبار القبلي.

ويمكن للمرء في التصميم 7 إجراء عدة مقارنات لتحديد تأثير المعالجةX التجريبية. فإذا كان وسط الاختبار البعدي للمجموعة (E) أكبر بدلالة، من وسط المجموعة الضابطة الأولى (C_1) وإذا كان وسط الاختبار البعدي (C_2) أكبر بدلالة، من نظيره (C_3)، فيكون لدينا دليل على فاعلية المعالجة التجريبية. ويمكن تحديد تأثير الظروف التجريبية في مجموعة أخذت الاختبار القبلي عن طريق مقارنة الاختبارات البعدية لكل من (E) و (C_1)، أو التغييرات القبلية والبعدية لكل من (E) و (C_1)، كما أن تأثير التجربة على المجموعة التي لم تأخذ الاختبار القبلي تتضح بمقارنة (C_2) و (C_3). وإذا كان متوسط الفروق بين درجات الاختبار البعدي ($E- C_1$) و (C_2 - C_3) متماثلاً تقريباً، فلابد أن للتجربة، عندئذ، تأثير متماثل تقريبا في المجموعات التي أخذت الاختبار القبلي والتي لم تأخذه.

يشمل التصميم 7 فعلاً إجراء تجربتين، واحدة مع اختبارات قبلية، وأخرى بدون اختبارات قبلية. فإذا اتفقت نتائج هاتين التجربتين كما أوضحنا أعلاه، فإن بوسع الباحث أن تكون له ثقة أكبر في الاستنتاجات.

ويكمن العيب الرئيسي لهذا التصميم، في الصعوبة التي ينطوي عليها تنفيذه في وضع عملي. فهناك حاجة لوقت وجهد أكثر لإجراء تجربتين متزامنتين، وهناك مشكلة تحديد العدد المتزايد لأفراد، نحتاجهم من النوع ذاته.

وثمة صعوبة أخرى تكمن في التحليل الإحصائي. فلا توجد أربع مجموعات كاملة من القياسات للمجموعات الأربعة. وكما أشرنا أعلاه، فبوسعنا إجراء مقارنات بين (E) و (C_1) وبين (C_2) و (C_3)، لكنه ليس هناك أي إجراء إحصائي واحد يمكنه الإفادة من القياسات الستة المتوفرة في الوقت ذاته. ويقترح (Campbell & Stanley, 1966) استخدام درجات الاختبار البعدي فقط عن طريق تصميم تحليل التباين ذي الطريقين. ويعتبر الاختبار القبلي متغيراً مستقلاً ثان إضافة إلى (X). ويكون التصميم كما يلي:

	بدون X	X
أخذت الاختبار القبلي	Y_2، ضابطة 1	Y_2 ، تجريبي
لم تأخذ الاختبار القبلي	Y_2، ضابطة 3	Y_2، ضابطة 2

ويحدد المرء من أوساط الأعمدة التأثير الرئيسي لـ (X). ويحدد من أوساط الصفوف التأثير الرئيسي للاختبار القبلي، ومن أوساط الخلية يحدد تفاعل الاختبار مع (X).

التصميمات العاملية Factorial Designs

كانت التصميمات المقدمة إلى الآن، عبارة عن تصميمات كلاسيكية ذات متغير واحد، يستخدم فيها الباحث متغيراً مستقلاً واحداً لإحداث تأثير على المتغير التابع. وعلى أي حال، ففي حالة الظواهر الاجتماعية المعقدة هناك، على العموم، عدة متغيرات تتفاعل في الوقت ذاته، وأية محاولة لتحديد الدراسة بمتغير واحد قد تفرض بساطة مصطنعة على وضع معقد. فالمتغير (X) وحده، قد لا يحدث التأثير نفسه كما يحدث في التفاعل مع (X) آخر، ومن ثم فإن استنتاجات من تصميمات ذات متغير واحد قد تكون دون معنى. فمثلاً، قد تعتمد فاعلية طريقة تدريس معينة على عدد من المتغيرات، كمستوى الاستعداد لدى الطلبة، وشخصية المدرس، والجو العام وما إلى ذلك. فالتدريس المدعوم بالحاسوب، مثلاً، قد يكون أكثر فاعلية مع الطلبة البطيئين منه مع الطلبة الأذكياء. إن تصميماً كلاسيكياً ذا متغير واحد لن يكشف عن التأثير التفاعلي للطريقة ومستوى الاستعداد. ويمكن زيادة حاصل المعلومات من التجربة، بشكل ملحوظ، عن طريق تأكيد التأثيرات المتزامنة لمتغيرين مستقلين أو أكثر في تصميم عاملي. لقد قيل، في الحقيقة، أن التقدم الفعلي في البحوث التربوية قد جاء مع تطوير فيشر للتصميمات العاملية.

إن التصميم العاملي، هو الذي يتم فيه تفعيل اثنين أو أكثر من المتغيرات في الوقت ذاته بغية دراسة التأثير المستقل لكل متغير في المتغير التابع وكذلك التأثيرات الناجمة عن التفاعلات بين متغيرات عدة.

والتصميمات العاملية هي على نوعين. في النوع الأول واحد فقط من المتغيرات المستقلة يتم استخدامه تجريبيا. وفي هذه الحالة يهتم الباحث أساسا بتأثير المتغير المستقل الواحد ولكن يجب أن يأخذ بنظر الاعتبار متغيرات أخرى قد تؤثر على المتغير التابع. وبصورة نموذجية، فإن هذه المتغيرات الأخرى هي متغيرات منسوبة / نوعية attribute كالجنس (gender)، والذكاء، والعرق، والوضع الاجتماعي الاقتصادي، والتحصيل وما شابه. ويمكن التحقق من تأثيرها (وضبطها في الوقت ذاته) عن طريق بناء المتغير التابع مباشرة في التصميم العاملي. ويقيم الباحث تأثير المتغير المستقل الرئيسي لكل مستوى من "المستويات" المتعددة لواحد أو أكثر من المتغيرات المنسوبة المستقلة. إن المستويات المختلفة للمتغير المنسوب تمثل بشكل نموذجي مجموعات الأشخاص المنتخبة بصورة طبيعية، كما تقوم مثلاً دراسة معينة باستخدام الطلبة الأذكياء والبطيئين لتحديد فاعلية أسلوب تدريس معين. إن بناء المتغيرات التابعة في تصميم عاملي لا يزيد من دقة التجربة وحسب بل كذلك ومن عموميتها. ولان المرء قادر على تحديد ما إذا كان للمعالجة أثار متماثلة تقريبا على كل المستويات، فإن تعميم الاستنتاجات التجريبية يزداد.

في النوع الثاني قد يجرى تفعيل جميع المتغيرات المستقلة تجريبيا. وهنا يهتم الباحث بعدة متغيرات مستقلة فعالة active ويرغب في تقييم كل من تأثيراتها المنفصلة والمدمجة. مثلاً، قد تقارن تجربة تأثيرات حجم الصف وكذلك إدخال التدريس المدعوم بالحاسوب على تعلم

المفاهيم العلمية. في هذه الدراسة سيفعّل كلا المتغيرين، وسيكون هناك مستويان للمتغير هما طريقة التدريس (أي المدعومة بالحاسوب مقابل التقليدية) وللمتغير الثاني، حجم الصف (أي كبير مقابل صغير). إن مثل هذا التصميم يسمح بتحليل الآثار الرئيسية لكل من المتغيرات التجريبية وتحليل التفاعل بين المعالجات.

التصميم 8: التصميم العاملي البسيط Design 8: Simple Factorial Design

لقد تم تطوير التصميمات العاملية على مستويات مختلفة من التعقيد. وابسط تصميم عاملي هو التصميم 2×2. وفي هذا التصميم يكون لكل متغير من المتغيرين المستقلين اثنان من المستويات.

التصميم 8: التصميم العاملي البسيط

المتغير 1 (X_2)	المتغير 2 (X_1) المعالجة أ	المعالجة ب
المستوى 1	الخلية 1	الخلية 3
المستوى 2	الخلية 2	الخلية 4

ولإيضاح ذلك، دعنا نفترض أن الباحث مهتم بمقارنة فاعلية نوعين من طرق التدريس – الطريقة أ والطريقة ب – حول تحصيل طلبة العلوم في الصف التاسع، مع الاعتقاد بأن هناك تأثيراً فارقاً لهاتين الطريقتين استنادا إلى مستوى الاستعداد العلمي للطلبة. ويصنف الباحث المجتمع الإحصائي إلى درجات استعداد عالية ومنخفضة، ويختار عشوائيا (60) فرداً من المجموعة العالية ويخصص (30) فرداً للطريقة أ، و (30) للطريقة ب. وتكرر هذه العملية للمجموعة منخفضة الاستعداد. ويخصص المدرسون عشوائيا لهذه المجموعات.

وفي تجربتنا الافتراضية، لدينا معالجتان تجريبيتان ومستويان من الاستعداد. ويبين الجدول 9.1 التصميم العاملي 2×2 لقياس تأثيرات طريقتي التدريس على تعلم الطلبة. لاحظ أن التصميم 2×2 يتطلب أربع مجموعات من الأفراد، حيث يتم توزيع عشوائي للأفراد، ضمن كل مستوى من مستوى الاستعداد، للمعالجتين.

وتمثل الدرجات في الخلايا الأربع، أوساط الدرجات للمجموعات الأربع حسب المتغير التابع وهو اختبار التحصيل العلمي. وإضافة إلى درجات الخلايا الأربع التي تمثل مختلف التوفيقات(*) (الترابطات) بين المعالجات والمستويات، هناك أربعة أوساط درجات هامشية، اثنان

(*) التوفيقة (Combination): مصطلح رياضي يشير إلى ربط عنصرين أو أكثر دون إعطاء أهمية للترتيب، خلافاً للتبديلة (Permulation)، فربط العناصر س، ص، ع ثنائياً يكون س ص، س ع، ص ع، حيث س ص مثل ص س وكذلك س ع مثل ع س ، وخارج الرياضيات تدعى توليفة. (المراجع)

للأعمدة واثنان للصفوف. وأوساط العمود الهامشية تخص الطريقتين أو المعالجتين، وأوساط الصفوف الهامشية تعود لمستويي الاستعداد.

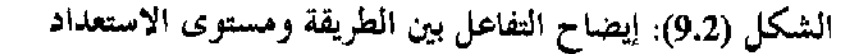
الجدول (9.1): مثال على التصميم العاملي

الاستعداد (X_2)	طريقة التدريس (X_1)		الوسط
	الطريقة أ	الطريقة ب	
عال	75.0	73.0	74
منخفض	60.0	64.0	62
الوسط	67.5	68.5	

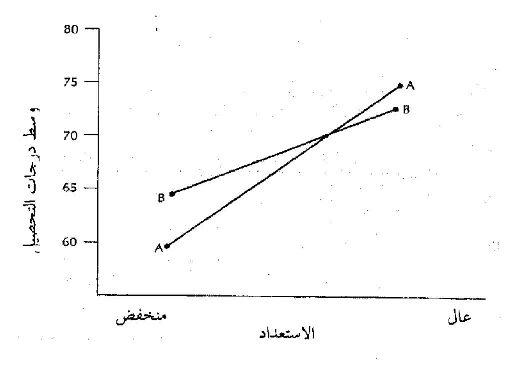
ومن البيانات المعطاة، نستطيع أن نحدد أولاً التأثيرات الرئيسية لمتغيرين مستقلين. وتشير درجات وسط المعالجة دون اعتبار لمستوى الاستعداد إلى التأثير الرئيسي للمعالجات. وإذا قارنا درجة الوسط لمجموعتي الطريقة أ وهي (67.5) مع درجة وسط مجموعتي الطريقة ب وهي (68.5) لوجدنا أن الفرق بين هذين الوسطين هو درجة واحدة فقط. وعليه فقد يغرينا الاستنتاج بان الطريقة المستخدمة لها تأثير قليل على المتغير التابع.

دعنا نتفحص أوساط درجات المستويات لغرض تحديد التأثير الرئيسي لـ X2 أو مستوى الاستعداد في درجات التحصيل. إن التأثير الرئيسي للمستويات لا يأخذ بنظر الاعتبار أي تأثير فارقي يعود إلى المعالجات. ودرجة الوسط لمجموعتي الاستعداد العالي هي (74) ودرجة الوسط لمجموعتي الاستعداد المنخفض هي (62) والفرق البالغ (12) درجة هو التأثير الذي يعزى إلى مستويي الاستعداد. فلمجموعتي الاستعداد العالي درجة وسط عالية بشكل ملحوظ، وهكذا، فبغض النظر عن المعالجة فإن أداء مجموعتي الاستعداد العالي افضل من مجموعتي الاستعداد المنخفض.

ويسمح التصميم العاملي للباحث بتقييم التفاعل بين المتغيرين المستقلين- أي التأثيرات المختلفة لأحدهما على مستويات مختلفة للآخر. وإذا كان هناك تفاعل معين فإن تأثير المعالجة على التعلم سيختلف بالنسبة إلى مستويي الاستعداد. وإذا لم يكن هناك تفاعل فإن تأثير المعالجة سيكون نفسه لكلا مستويي الاستعداد. ومن فحص الجدول 9.1، يمكننا أن نرى أن وسط الطريقة أ أعلى من وسط الطريقة ب لمجموعة الاستعداد العالي. وهكذا فإن توفيقات معينة للمعالجة ومستوى الاستعداد تتفاعل لتوفير لإعطاء مكاسب أكبر من توفيقات أخرى. ويتضح تأثير هذا التفاعل بين الطريقة ومستويات الاستعداد بشكل بياني في الشكل 9.2. فإذا كان هذا التفاعل دالاً إحصائيا فإنه يمكننا أن نستنتج أن فاعلية الطريقة تعتمد على الاستعداد. فالطريقة أ هي أكثر تأثيرا مع طلبة الاستعداد العالي والطريقة ب هي أكثر تأثيرا مع مجموعة الاستعداد المنخفض.



الشكل (9.2): إيضاح التفاعل بين الطريقة ومستوى الاستعداد



دعنا نفحص مجموعة أخرى من البيانات المتوفرة في الدراسة العاملية الافتراضية 2×2. يبين الجدول 9.2 نتائج الدراسة المصممة لبحث تأثير طريقتي تدريس على التحصيل. ومرة أخرى، نظراً لأن الباحث يتوقع أن الطريقة قد تكون مؤثرة بصورة متباينة استنادا إلى استعداد الفرد فإن الخطوة الأولى هي التمييز بين مستويين من الاستعداد. فالأفراد ضمن كل مستوى يخصصون بشكل عشوائي إلى طريقتين.

الجدول 9.2: مثال على التصميم العاملي

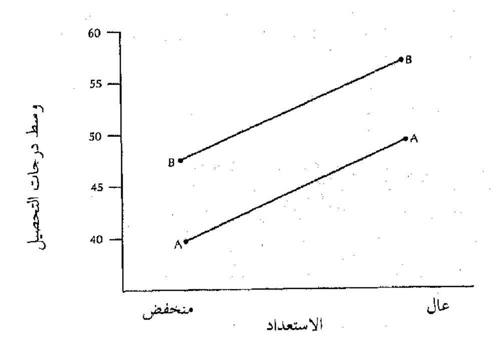
الاستعداد (X_2)	المعالجة (X_1) الطريقة أ	الطريقة ب	الوسط
عال	50	58	54
منخفض	40	48	44
الوسط	45	53	

وعقب التجربة، تجرى اختبارات التحصيل ويتم تسجيل الدرجات لكل فرد. فإذا قارنا درجة الوسط للمجموعتين التي تم تدريسهما حسب الطريقة ب (53) مع تلك الخاصة بالمجموعتين اللتين تم تدريسهما حسب الطريقة أ (45) لرأينا أن الأولى هي أعلى نوعا ما. وعليه، تبدو الطريقة ب أكثر تأثيرا من الطريقة أ. إن الفرق بين وسطي مستويي الاستعداد على التأثيرات الرئيسية للاستعداد هو (54-44)10. وبغض النظر عن المعالجة، فإن أداء مجموعة

الاستعداد العالي أفضل من مجموعة الاستعداد المنخفض. ولا تكشف البيانات عن وجود تفاعل بين المعالجة والمستويات. وتبدو الطريقة ب أكثر فاعلية بغض النظر عن مستوى الاستعداد. وبعبارة أخرى، فإن المعالجات والمستويات مستقلة عن بعضها. فالافتقار للتفاعل يتضح بيانيا في الشكل 9.3. وليس ممكناً إيضاح وجود أو غياب مثل هذا التفاعل دون استخدام التصميم العاملي.

ويمكن توسيع التصميم العاملي إلى تجارب أكثر تعقيدا يوجد فيها عدد من المتغيرات المستقلة، وتشير القيم العددية لهذه الأرقام إلى عدد المستويات الخاصة بالمتغيرات المستقلة المحددة. فمثلاً، في التصميم العاملي 4×3 ×2 هناك ثلاثة متغيرات مستقلة مع مستويين، وثلاثة، وأربعة على التعاقب. وقد تستخدم مثل هذه التجربة طريقتي تدريس وثلاث مستويات قدرة وأربعة صفوف دراسية. ونظرياً، قد يشمل التصميم العاملي أي عدد من المتغيرات المستقلة مع أي عدد من المستويات لكل منها. وعلى أية حال، عندما تستخدم عدة عوامل، أو تضبط تزامنياً، فإن الدراسة والتحليل الإحصائي يصبحان غير عمليين وقد تكون بعض التوفيقات مصطنعة. إن عدد المجموعات المطلوبة للتصميم العاملي هو حاصل ضرب القيم الرقمية التي توضح التصميم العاملي. ففي تصميم 4 × 3 × 2 ستكون هناك حاجة إلى 24 مجموعة بغية تمثيل جميع التوفيقات لمستويات المتغيرات المستقلة المتعددة المختلفة. إن مجرد فكرة التعقيدات المترتبة على تنظيم أعداد كبيرة من الأفراد تحت أعداد كبيرة من الظروف، ربما سيساعد القارئ على فهم سبب محاولة، اغلب البحوث التربوية، الإجابة على أسئلتها مع ابسط التصميمات المحتملة، حتى وإن كان بالإمكان معالجة التحليل الإحصائي بسهولة عن طريق الحواسيب الإلكترونية.

الشكل (9.3): إيضاح افتقار التفاعل بين الطريقة ومستوى الاستعداد



وتكمن مزايا التصميم العاملي بإنجازه في تجربة واحدة ما قد يحتاج، خلافاً لذلك، إلى دراستين منفصلتين أو أكثر، ويوفر فرصة لدراسة التفاعلات التي تعتبر مهمة دائما في البحوث التربوية، كما يؤمن اختبارا أكثر قوة للفرضيات.

التصميمات شبه التجريبية Quasi - Experimental Designs

يفضل الباحثون التصميمات التي تؤمن ضبطا تجريبيا كاملا من خلال استخدام إجراءات العشوائية. وهذه هي التصميمات التجريبية الحقيقية كما قدمت في الجزء السابق (التصميمات من 3 إلى 8). وهناك مواقف عديدة في البحوث التربوية التي لا يمكن فيها إجراء تجربة حقيقية. فلا الضبط التام على الجدولة الزمنية للظروف التجريبية، ولا القدرة على العشوائية يمكن تحقيقها دائما. مثلاً، في البحث الذي أجرى في وضع صفي، قد لا يكون بالإمكان للباحث تخصيص الأفراد عشوائيا في مجموعات. فيجب على المرء، في هذه الحالة، استخدام التصميمات التي تؤمن أكبر ضبط ممكن في ظل الوضع القائم. وتدعى التصميمات التي لا تشمل التخصيص العشوائي بالتصميمات شبه التجريبية، وتستخدم عندما لا تكون التصميمات التجريبية الحقيقية عملية. ونظراً لأن التصميم شبه التجريبي لا يؤمن ضبطا كاملاً، فإن من المهم جدا أن يدرك الباحث الأخطار على كل من الصدقين الداخلي والخارجي، ويأخذ بنظر الاعتبار هذه العوامل في التفسير. وعلى أية حال، أوضح (Campbel & Stanley) أن الدراسات شبه التجريبية "تستحق الاستخدام عندما يتعذر التحقيق الدقيق (1966, p. 205). وتسمح هذه التصميمات للمرء بالتوصل إلى استنتاجات معقولة حتى إذا كان الضبط الكامل غير ممكن.

التصميم 9: تصميم المجموعة الضابطة غير العشوائية ذو الاختبارين قبلي وبعدي
Design 9: Nonrandomized Control Group, Pretest - Postest Design

رغم أن التخصيص العشوائي للأفراد في مجموعات هو الأمر المثالي، فإنه ليس ممكنا من حيث الممارسة. ففي وضع مدرسي نموذجي، لا يمكن تعطيل جداول الدراسة ولا الصفوف التي جرى تنظيمها، بغية تنفيذ دراسة البحث. ففي مثل هذه الحالة، يكون من الضروري استخدام مجموعات حسب وضعها، لأنها منظمة في صفوف، أو في مجموعات تامة أخرى.

التصميم 9: تصميم المجموعة الضابطة غير العشوائية مع اختبارين قبلي وبعدي

المجموعة	الاختبار القبلي	المتغير المستقل	الاختبار البعدي
E	Y_1	X	Y_2
C	Y_1	–	Y_2

قد يسمح للباحث إجراء تجربة مع أربعة صفوف في اللغة الإنجليزية من طلبة السنة الأولى

في مدرسة ثانوية. ولان الصفوف تلتقي في أوقات مختلفة، فإنه لا يمكن توزيع الأفراد عشوائياً في المعالجات. وعلى أية حال، يمكن للباحث استخدام إجراء عشوائي لتحديد أي صفين سيكونان تجريبيين وأي اثنين ضابطين. فجميع الأفراد، يأخذون الاختبار القبلي قبل بدء التجربة، والاختبار البعدي في نهايتها. وحيث أن المجموعات التجريبية والضابطة تأخذ ذات الاختبار القبلي والبعدي، وان التجربة تستغرق ذات الزمن بالنسبة إلى جميع الأفراد، فإن إجراء الاختبار، واستخدام الأداة والنضوج والتسرب، ليست مشكلات للصدق الداخلي.

فإذا كان الباحث يتحكم تماماً بالمعالجة فإن التاريخ لا يعد مشكلة. وإذا قام الباحث بمجرد الإشراف على المدرسين المنتظمين الذين يوصلون المعالجات التجريبية والضابطة، فإن الفروقات بين المدرسين قد تؤثر على النتائج بشكل منتظم.

يستخدم كل من التصميم 9 والتصميم 5 – تجربة المجموعة الضابطة حسب توزيع عشوائي للأفراد واختبارين قبلي وبعدي – الفروقات بين الاختبار القبلي والاختبار البعدي كمتغير تابع. إن اللاعشوائية في التصميم 9 مسؤولة عن ثلاثة تهديدات للصدق الداخلي، حيث تتجنبها العشوائية في التصميم5:

1– رغم أن الاختيار ذاته لا يعد مشكلة صدق داخلي، ونظراً لأنه لم يجر تخصيص الأفراد من أجل هدف التجربة، وأن المجموعات التي تحظى بالمعالجات التجريبية والضابطة قد اختيرت بشكل عشوائي، فإن تفاعل الاختيار والنضوج قد يكون مشكلة صدق داخلي خطيرة. فلنفترض أن الشعبة الأولى في مادة الإنجليزية تلتقي في ذات الساعة التي يلتقي بها صف علاج (تقوية) في الرياضيات بينما الشعبة الثالثة تلتقي في ذات الساعة التي يلتقي بها صف الجبر المتقدم. وهنا نتوقع أن الشعبة الأولى كمجموعة ستحوز على أكبر فائدة في الأداء الأكاديمي، لأن العديد من الطلبة الضعفاء هم في درس الرياضيات العلاجي في ذات الساعة. فيتوقع من الشعبة الثالثة أن تبدي أقل فائدة، لأن العديد من افضل الطلبة هم في صف الجبر المتقدم في الوقت ذاته. إن توقيت دروس الرياضيات يؤثر على بنية صفوف الإنجليزية، وهذا يؤثر على الكسب المتوقع بسبب التفاعل بين الاختيار والنضوج.

ويمكن أن يكون تفاعل الاختيار والنضوج مشكلة صعبة، بشكل خاص، عندما يقارن المتطوعون مع غير المتطوعين. فمثلاً، هناك برنامج تحسين القراءة يقدم بعد المدرسة للراغبين فيه. ولم تبين أوساط الاختبار القبلي لاختبار القراءة أي فرق بين من تطوعوا لبرنامج ما بعد المدرسة ومن لم يتطوعوا له. فإذا كانت درجات ما بعد المعالجة قد كشفت عن كسب / فائدة لمجموعة المعالجة أكبر من المجموعة الضابطة، فإنه لا يمكننا أن نعزو الكسب الأكبر إلى المعالجة بشكل أكيد. فمن المحتمل أن من رغبوا في المساهمة في برنامج ما بعد المدرسة، كانوا معنيين بقراءتهم، وعليه يحتمل لهم أن يحوزوا على كسب اكبر في القراءة سواء حصلوا على المعالجة أم لم يحصلوا عليها.

2- يعد الارتداد الإحصائي مشكلة صدق محتملة أخرى للتصميم 9. ويشير هذا المصطلح إلى ميل الدرجات المتطرفة إلى الارتداد (التحرك) نحو الوسط المألوف في القياس التالي. إن تأثير مثل هذا الارتداد يمكن إدخاله في التصميم، إذا سحبت المجموعات المستخدمة من مجتمعات إحصائية لها أوساط مختلفة. ورغم أن المجموعات تكون متكافئة في الاختبار القبلي، إلا أن تأثير الارتداد الذي يحدث يمكن أن يؤدي إلى تغير من الاختبار القبلي إلى الاختبار البعدي الذي يفسر، خطأ، على انه تأثير تجريبي.

دعنا نفترض أن للمجموعة التجريبية في الدراسة وسط مقداره (75) في اختبار قبلي أدنى من وسط مجتمعها الأصلي، بينما للمجموعة الضابطة وسط اختبار قبلي مقداره (75) أعلى من وسط مجتمعها. ونظراً لأن كل مجموعة سوف ترتد نحو وسط المجتمع الأصلي عند إعادة الاختبار، فيتوقع للمجموعة التجريبية أن يكون لها وسط أعلى في الاختبار البعدي سواء تم إدخال (X) أم لا، ومن ناحية أخرى سيرتد وسط المجموعة الضابطة إلى الأسفل. وستبدو المجموعة التجريبية كأنها تحرز تقدما أكثر خلال مسار الدراسة من المجموعة الضابطة، وهو ما يعزى على الأغلب، وبشكل خاطئ، إلى تأثير (X).

3- تبرز مع التصميم (9) مشكلات خطيرة تخص الصدق الداخلي مع التغير أو "الكسب" في الدرجات لأن الأفراد لم يوزعوا بصورة عشوائية بمجموعات المعالجة أو الضبط. ومع ذلك فإنه لا يمكن دراسة مشكلات تغير الدرجات بالتفصيل في هذا الكتاب، دعنا الآن نشير إلى بعض هذه الصعوبات. فهناك عادة ارتباط سالب بين درجات الاختبار القبلي والكسب الحاصل من اختبار قبلي إلى اختبار بعدي. فهل يعني هذا أن الطلبة ذوي الدرجات المنخفضة في البداية، يتعلمون أكثر (حسب قياس درجات التغير(*)) من الطلبة ذوي الدرجات العالية في البداية؟ ربما لا. فيحتمل للارتباط السالب أن يعزى إلى خواص القياس النفسي المتميزة لدرجات التغير. فمن ناحية، يكون لأغلب الاختبارات التربوية سقف، بما يعني أن مدى التحصيل في بنود الاختبار محدود. فإذا أجاب تلميذ على (92) بنداً بشكل صحيح في اختبار قبلي من 100 بند، فيمكن، فقط، لهذا التلميذ أن يكسب (8) نقاط في الاختبار البعدي. ومن ناحية أخرى، يمكن لتلميذ درجاته (42) في الاختبار القبلي، أن يكسب (58) نقطة / درجة. وبسبب تأثير السقف هذا، فإن التلاميذ في مجموعة التحصيل العالي في الاختبار القبلي، سيكونون مقيدين بدرجة تغير منخفضة في الاختبار البعدي.

إن المشكلات المرتبطة بدرجات التغير هي مجرد مخاطر للصدق الداخلي عندما لا يتم توزيع الأفراد بشكل عشوائي في المعالجة. فمع العشوائية، سيكون لأي تحريف / تشويه نتيجة

(*) درجة التغير (change score): تشير إلى الفرق بين درجتي الفرد في المتغير التابع حسب الاختبارين القبلي والبعدي- (المراجع)

استخدام درجات متغيرة، على المدى البعيد، ذات التأثير على كل من المجموعتين التجريبية والضابطة. وعند استخدام التصميم (9) تكون المشكلات الخاصة بدرجات التغير خطيرة، بصورة خاصة، عندما يختلف وسطا درجات الاختبار القبلي للمجموعتين التجريبية والضابطة بشكل كبير لأن تأثير السقف والارتداد قد يؤثر على درجات الاختبار البعدي لمجموعة واحدة أكثر من الأخرى. فمثلاً، إذا استخدم التصميم (9) لمقارنة آثار طريقتين في تدريس التهجئة تستخدم فيها أشكال متكافئة في اختبار تهجئة من مئة بند كاختبار قبلي وبعدي، وكان لمجموعة منهما، وسط اختبار قبلي يبلغ (80) والأخرى (50)، فإن تأثير السقف سيحدّ من الكسب الممكن للأولى أكثر من الثانية.

إن الأخطار على الصدق الداخلي في التصميم (9) مماثلة للأخطار الموجودة في التصميم (5). وعلى أية حال، فإن مزية التصميم (9) تكمن في إمكانية ضبط التأثيرات التفاعلية(*) للتجريب، بسهولة أكثر مما في التصميم (5). وعند استخدام الصفوف (التامة) قد يكون الأفراد أقل إدراكا لإجراء تجربة مما عليه الحال، عندما يسحب الأفراد من صفوف ويوضعون في حصص تجريبية. وهذا يساهم في تعميم الاستنتاجات. وبالمناسبة، قد يمكن ملاحظة أن للباحث، في وضع مدرسي، فرصة أكبر للحصول على موافقة إدارية لإجراء التجربة، إذا استخدمت صفوف تامة كما في التصميم (9).

وكلما كانت المجموعتان التجريبية والضابطة متماثلتين في بداية التجربة، وازداد تأكيد هذا التماثل بأوساط متماثلة في الاختبار القبلي للمجموعتين، كانت نتائج دراسة المجموعة الضابطة غير العشوائية في الاختبارين القبلي والبعدي أكثر صدقا. وإذا كانت درجات الاختبار القبلي متماثلة وأمكن إيضاح أن التفاعل بين الاختيار والنضوج والارتداد لا يحتمل لها أن تفسر فروقات الاختبار البعدي، فإن نتائج هذا التصميم شبه التجريبي صادقة تماماً.

وحتى إذا كان وسطا المجموعتين مختلفين بشكل ملحوظ قبل المعالجة، فإن مجموعة ضابطة غير مكافئة هي أفضل من عدم وجودها على الإطلاق. إن هذا التصميم خيار أفضل من التصميمات قبل التجريبية التي ليس لها مجموعة ضابطة. ويمكن توسيع التصميم (9) لاستخدام أكثر من مجموعتين.

التصميم 10: تصميم الموازنة (تدوير المجموعات)

Design 10: Counterbalanced Design

التصميم (10)، هو تصميم آخر يمكن استخدامه مع مجموعات صفية تامة وتدويرها في فترات

(*) التأثير التفاعلي / الراجع (reactive effect): مصطلح يشير إلى تأثير البحث على أفراد العينة، وتأثيرهم بالمقابل على نتائج البحث. (المراجع)

خلال التجربة. فمثلاً، قد تستخدم المجموعتان (1) و (2) الطريقتين أ و ب على التعاقب، بالنسبة للنصف الأول من التجربة، ثم يتم تغيير الطريقتين خلال النصف الثاني. إن الجانب المميز للتصميم (10) هو أن جميع الأفراد يحظون بجميع المعالجات التجريبية في وقت ما خلال التجربة. وبذلك فإن التصميم ينطوي على سلسلة تدويرات، تنجم عن تكرارات متطابقة، حيث يتم في كل تكرار تعديل المجموعات بحيث تتعرض كل مجموعة لكل (X) في نهاية التجربة. ويختلف سياق التعرض للوضع التجريبي بالنسبة لكل مجموعة. وعادة ما يستخدم تصميم الموازنة عندما يراد اختبار عدة معالجات، ولكنه قد يستخدم مع معالجتين فقط.

التصميم 10: نموذج تصميم موازنة

تكرار / تدوير	المعالجات التجريبية			
	X_1	X_2	X_3	X_4
1	المجموعة أ	ب	ج	د
2	المجموعة ج	أ	د	ب
3	المجموعة ب	د	أ	ج
4	المجموعة هـ	ج	ب	أ
	متوسط العمود	متوسط العمود	متوسط العمود	متوسط العمود

إن كل صف في التصميم (10) يمثل تدويراً واحداً. وفي كل تدوير يتم تبديل المجموعات بحيث أن المجموعة أ تجرب أولاً X_1 ثم X_2 ثم X_3 وأخيراً X_4. بينما تحتوي كل خلية في التصميم وسط الدرجات في المتغير التابع للمجموعة والمعالجة والتدوير / التكرار الموضح. ويوضح وسط الدرجات لكل عمود أداء المجموعات الأربعة في المتغير التابع تحت المعالجة الممثلة بالعمود.

ويمكن للمدرس في الصف أن يستخدم دراسة الموازنة لمقارنة فاعلية طريقتي التدريس على التعلم في مجال العلوم. فبوسع المدرس اختيار صفين ووحدتي علوم متضاهيين، أي يتماثلان تقريباً، من حيث الصعوبة والطول وما إلى ذلك. ومن المهم أن تكون الوحدات متكافئة في التعقيد وصعوبة المفاهيم المعنية. وخلال التدوير الأول للتصميم يتم تدريس الصف 1 الوحدة 1 حسب الطريقة (أ) والصف 2 حسب الطريقة (ب). ويجري اختبار تحصيل للوحدة 1 لكلا المجموعتين. ثم يتم تدريس الصف 1 الوحدة 2 حسب الطريقة (ب) والصف 2 حسب الطريقة (أ). ثم يتم اختبار كليهما في الوحدة 2. إن هذا الترتيب مبين في الجدول 9.3.

الجدول 9.3: مثال عن التصميم المتقابل التوازن

التكرار	المعالجات التجريبية	
	الطريقة أ	الطريقة ب
(الوحدة) 1	الصف 1	الصف 2
(الوحدة) 2	الصف 2	الصف 1
	وسط العمود	وسط العمود

وبعد الدراسة، يتم حساب وسطي العمودين لإيضاح وسط التحصيل لكلا المجموعتين (الصفين) حين تدريسهما حسب الطريقة الموضحة عن طريق عنوان العمود. وتوضح مقارنة درجات وسطي العمودين، باستخدام تحليل التباين، فاعلية الطريقتين على التحصيل في العلوم.

يتغلب التصميم (10) على بعض درجات الضعف في التصميم (9)، أي عندما يتحتم استخدام صفوف تامة، فإن الموازنة تعطي الفرصة لتدوير أية فروقات قد توجد بين المجموعتين. وحيث أن المعالجات تجري على كل المجموعات، فإن النتائج الحاصلة لكل (X) لا يمكن أن تعزى إلى فروقات موجودة مسبقا لدى الأفراد. فإذا كانت مجموعة أكثر ذكاءاً، في المتوسط، من المجموعة الأخرى فإن كل معالجة (X) ستستفيد من الذكاء الأعلى.

العيب الرئيسي للتصميم (10)، هو أنه قد يكون هناك احتفاظ بتأثير من (X) إلى آخر. وعليه ينبغي استخدامه عندما لا تكون للمعالجات التجريبية، في تعرض لمعالجة واحدة، تأثير على المعالجات التالية. وقد يكون هذا الشرط صعب التحقيق في الكثير من البحوث التربوية. علاوة على ذلك، هناك ضرورة لإقامة تكافؤ في مادة التعلم المستخدمة في التدويرات المختلفة. وقد لا يكون من الممكن دائما تحديد وحدات مادة متكافئة. وثمة عيب آخر لهذا التصميم، هو احتمال أن يصاب الطلبة بالملل من الاختبارات المتكررة التي تتطلبها هذه الطريقة.

تصميمات السلسلة الزمنية Time Series Designs

التصميم 11: تصميم السلسلة الزمنية ذو المجموعة الواحدة

Design 11: One -Group Time Series Design

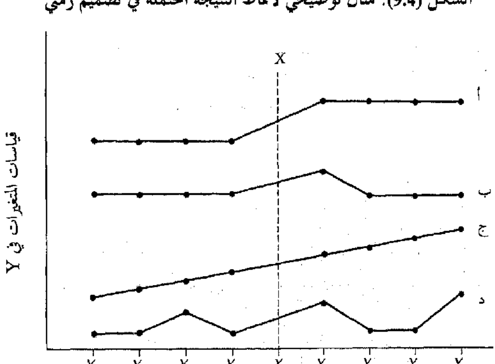
ينطوي التصميم 11 على قياس دوري لمجموعة واحدة وإدخال معالجة تجريبية في هذه السلسلة الزمنية للقياسات. وكما يوضح التصميم فإنه يتم اخذ القياسات في متغير تابع (Y)، ثم يتم إدخال (X) وتجرى قياسات إضافية في Y كذلك..

التصميم 11: تصميم السلسلة الزمنية ذو المجموعة الواحدة

Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8

ويمكن بمقارنة القياسات قبل وبعد (X) تقدير تأثير X على أداء المجموعة في (Y). ويمكن استخدام تصميم السلسلة الزمنية، في وضع مدرسي، لدراسة تأثيرات تغير كبير في السياسة الإدارية على الحوادث الانضباطية. وقد يتضمن دراسة قياسات متكررة لاتجاهات الطلبة والتأثير الحاصل نتيجة إدخال فيلم وثائقي مصمم لتغيير الاتجاهات.

الشكل (9.4): مثال توضيحي لأنماط النتيجة المحتملة في تصميم زمني



ويوضح الشكل 9.4 بعض الأنماط المحتملة من دراسات السلسلة الزمنية التي أدخلت إليها معالجة تجريبية. وهو يبين سلسلة قياسات (Y_1) حتى (Y_8) مع إدخال المعالجة التجريبية عند نقطة (X). ويمكننا أن نقيم تأثير (X) عن طريق تفحص ثبات / استقرار القياسات المتكررة.

ومن تفحص الفرق بين (Y_4) و (Y_5) في النمط (أ)، ربما يجد المرء ما يبرر الافتراض أن (X) تؤثر على المتغير التابع. ويشير النمط (ب) إلى احتمال التأثير التجريبي المؤقت لـ (X).

وعلى أية حال، لا يمكن للمرء افتراض أن (X) يحدث التغيير في أي من النمط (ج) أو النمط (د). ويبدو أن النمط (ج) ناتج عن النضج أو عن تأثير مماثل. أما الطبيعة غير المنتظمة للنمط (د) فتشير إلى عمل العوامل الدخيلة.

ويشابه التصميم 11، التصميم 1 بكونه يستخدم القياسات القبلية والبعدية ويفتقر إلى المجموعة الضابطة. وعلى أية حال، فإن له بعض المزايا على التصميم 1، بما يجعله أكثر فائدة في البحوث التربوية. فالاختبار المتكرر يوفر ضبطاً لبعض المخاطر المألوفة على الصدق الداخلي. فالنضوج، والاختبارات، والارتداد يمكن استبعادها كتفسيرات مقبولة في ظاهرها للتغير الذي يحدث بين (Y_4) و (Y_5)، إذا لم تحدث مثل هذه التغيرات في الفترات الزمنية السابقة قيد الملاحظة. ويوصى بعدم إجراء أي تغيير في أدوات القياس خلال فترة الدراسة الزمنية. وبهذه الطريقة يستبعد المرء التغييرات في الأداة كتفسير محتمل للفرق (Y_5-Y_4).

أما الضعف الكبير للتصميم 11 فهو إخفاقه في ضبط التاريخ، أي أنه لا يمكن للمرء استبعاد إمكانة انه ليس (X)، بل حادث آني ما، قد احدث التغيير الملحوظ. وربما يعزى التغير إلى عوامل مثل التغييرات الموسمية أو الجوية أو المدرسية مثل الامتحانات. ففي دراسة صممت لتقييم تأثير معالجة فيلم محاضرة عن اتجاهات الطلبة نحو الأقليات يطرح السؤال التالي: إلى أي مدى تتأثر فيه قياسات الاتجاهات بشغب أقلية في مدينة بعيدة، تناقلته وسائل الإعلام على صعيد قومي؟ إن المدى الذي يعد فيه التاريخ (أحداث معاصرة خارج نطاق السيطرة) عاملا تفسيريا مقبولاً يجب أن يأخذه الباحثون بنظر الاعتبار حين يحاولون تفسير استنتاجاتهم.

ويجب كذلك أن يأخذ المرء بنظر الاعتبار الصدق الخارجي للتصميم الزمني. ونظراً لوجود اختبارات متكررة، فربما يكون هناك نوع من تأثير التفاعل للاختبار الذي قد يقيّد الاستنتاجات بمجتمعات الأفراد المعرضين للاختبار المتكرر. وعلى أية حال، طالما أن القياسات هي ذات نوع نمطي روتيني مستخدم في الأوضاع المدرسية، فإن ذلك لا يمكن أن يعتبر تقييداً خطيراً. وعلاوة على ذلك، قد يحدث تفاعل اختيار – X، خصوصا إذا اختار أحدهم مجموعة معينة قد لا تكون نموذجية.

ويمكن للتفسير الإحصائي أن يكون مشكلة خاصة مع بيانات زمنية. والاختبارات الاعتيادية للدلالة قد لا تكون مناسبة مع التصميم الزمني. ويمكن للقارئ الرجوع إلى كتاب (Campbell & Stanley, 1966) لمناقشة الاختبارات الإحصائية التي قد تستخدم مع هذا التصميم.

التصميم 12: تصميم السلسلة الزمنية ذو المجموعة الضابطة

Design 12: Control Group Time Series Design

يعد التصميم 12 امتدادا للتصميم 11 ليشمل المجموعة الضابطة. والمجموعة الضابطة التي تمثل مرة أخرى صفاً تاماً، قد تقاس في الوقت ذاته كالمجموعة التجريبية (E) لكنها لن تجرب المعالجة (X). ويتغلب هذا التصميم على ضعف التصميم 11 – أي الإخفاق في ضبط التاريخ كمصدر

لتباين دخيل. وتسمح المجموعة الضابطة بالمقارنة الضرورية. فإذا أبدت المجموعة (E) كسبا من Y_4 حتى Y_5 ولم تبد المجموعة الضابطة (C) أي كسب، فلابد للتأثير أن يعزى إلى X وليس إلى أحداث معاصرة، التي ربما أثرت على كلا المجموعتين.

وتشمل التباينات الأخرى في تصميم السلسلة الزمنية على إضافة مجموعات ضابطة أكثر، وملاحظات أكثر، ومعالجات تجريبية أكثر.

التصميم 12: تصميم السلسلة الزمنية ذو المجموعة الضابطة

المجموعـــة									
E	Y_1	Y_2	Y_3	Y_4	X	Y_5	Y_6	Y_7	Y_8
C	Y_1	Y_2	Y_3	Y_4	–	Y_5	Y_6	Y_7	Y_8

مشكلات الصدق مع التصميمات التجريبية
Validity Problems With Experimental Designs

يلخص الجدول 9.4 بعض مصادر انعدام الصدق في التصميمات التجريبية ذات المتغير الواحد. ولا يمكن اعتماد هذه الخلاصة الموجزة كدليل وحيد في اختيار تصميم معين. بل يجب أن يصاحب ذلك دراسة عميقة للتقديم المؤهل الذي يظهر في النص، بما يتيح للقارئ أن يفهم عوامل القوة والضعف الخاصة بكل تصميم.

الجدول 9.4 العوامل التي تهدد الصدق الداخلي للتصميمات التجريبية

التصميمات (*)												
Quasi-experimental				True Experimental						Pre-experimental		
12	11	10	9	8	7	6	5	4	3	2	1	مصادر الضعف
+	–	+	+	+	+	+	+	+	+	+	–	التاريخ المعاصر **
+	+	+	+	+	+	+	+	+	+	؟	–	عمليات النضج
+	+	+	+	+	+	+	+	+	+	+	–	أساليب الاختبار القبلي
+	؟	+	+	+	+	+	+	+	+	+	–	أدوات القياس
+	+	+	؟	+	+	+	+	+	+	+	؟	الارتداد الإحصائي
+	+	+	+	+	+	+	+	+	+	–	+	الاختيار التمييزي / الفارقي للأفراد
+	+	+	+	+	+	+	+	+	+	–	+	التسرب التجريبي
+	+	؟	+	+	+	+	+	+	+	–	–	التفاعل بين الاختبار والنضج وما أشبه
–	–	+	–	–	–	–	–	–	–	–	–	اتجاهات الأفراد
–	–	–	–	–	–	–	–	–	–	–	–	التنفيذ

* **التصميمات كما يلي:**

1- مجموعة واحدة واختباران قبلي وبعدي.
2- مقارنة مجموعة ثابتة / مستقرة.
3- مجموعة ضابطة حسب تخصيص عشوائي للأفراد مع اختبار بعدي فقط.
4- مجموعة ضابطة حسب تناظر عشوائي للأفراد مع اختبار بعدي فقط.
5- مجموعة ضابطة حسب تخصيص عشوائي مع اختبارين قبلي وبعدي.
6- تصميم سولومون ذو المجموعات الثلاث.
7- تصميم سولومون ذو المجموعات الأربع.
8- العاملي البسيط.
9- مجموعة ضابطة غير عشوائية مع اختبارين قبلي وبعدي.
10- الموازنة / التدوير.
11- سلسلة زمنية ذات مجموعة واحدة
12- سلسلة زمنية ذات مجموعة ضابطة

** إن علامة الزائد تشير إلى ضبط العامل، بينما تشير علامة الناقص، إلى الافتقار للضبط، وعلامة الاستفهام تشير إلى مصدر يثير القلق.

التصميمات التجريبية ذات الفرد الواحد

Single-Subject Experimental Designs

يبدو أن التصميم التجريبي ذي الفرد الواحد يمثل تناقضا في الاصطلاحات. فكيف يمكن لتجربة أن تجرى مع عينة حجمها واحد فقط؟ ومن الواضح، لا يمكن أن يكون هناك تخصيص عشوائي أو استخدام للمجموعات الضابطة. ومع ذلك، فإن البحوث التي تنطوي على فرد واحد، قد أصبحت مشهورة على مدى الـ 25 سنة الماضية. ويدافع مناصرو هذه المنهجية في القول بأنه يمكن للضبط التجريبي أن يتحقق بغير الطرق التقليدية. فبعد وصف الطريقتين الرئيسيتين للبحث ذي الفرد الواحد، والأساس المنطقي له سوف نتفحص نقاط القوة وقيود هذا النوع من البحوث، مقارنة بالتصميمات الأخرى الأكثر تقليدية.

لقد كان لدراسة الفرد، دائماً، مكان في البحوث التربوية والنفسية. فدراسات فرويد للحالة، وملاحظات بياجيه عن الأطفال هي أمثلة بارزة على ذلك. ورغم أن دراسات الحالة (انظر الفصل 11) وتجارب الفرد الواحد تدرس كلاهما الفرد، حيث يفعّل الباحث، بشكل مدروس، في تجارب الفرد الواحد متغيراً مستقلاً واحداً أو أكثر، بينما يلاحظ الباحث في دراسة الحالة تفاعل الأفراد مع الأحداث التي تقع بصورة طبيعية.

لقد كانت تصميمات الحالة الواحدة مفيدة، بشكل خاص، في التطبيقات السريرية حيث يجري التركيز على القيمة العلاجية للتدخل بالنسبة للمريض. فقد يرغب مدرس أطفال، يعانون من العوق الحاد، مثلاً في الحصول على معلومات تخص فاعلية إجراء محدد مع طفل واحد. لقد شعر بعضهم أن الدراسات التي تذكر وسط أو متوسط الفروقات بالنسبة للمجموعات قد تكون محدودة الفائدة عند معالجة شخص محدد.

تعتبر التصميمات ذات الفرد الواحد امتدادات، بشكل خاص، لتصميم السلسلة الزمنية ذي المجموعة الواحدة، شبه التجريبي (التصميم 11). وأكثر التصميمات شهرة هما تصميم (ABAB) وتصميم خط القاعدة المضاعف.

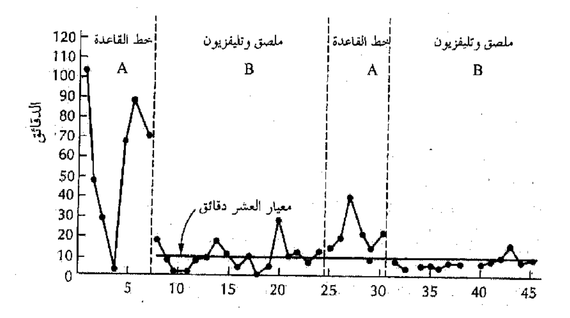
تصميمات ABAB Designs ABAB

يتكون تصميم (ABAB) من فترة بلا معالجة أو خط القاعدة(*) (A) حيث يقاس خلالها السلوك المعني بصورة متكررة. وأمثلة أنواع السلوك هذا، هي عدد المرات التي تترك فيها طالبة ذات عجز إدراكي / عقلي مقعدها، أو عدد المرات التي يضرب فيها طفل متوحّد / انعزالي رأسه. إن تقييم ما قبل المعالجة يفيد كفترة ضابطة تتم مقارنتها بتأثيرات المعالجة. وبعد إقامة صورة ثابتة لسلوك ما قبل المعالجة تتم المبادرة بالمرحلة (B) أو المعالجة. وفي مرحلة المعالجة، قد يعطى الطفل قطعة رمزية (يمكن تبادلها مع امتيازات مرغوبة) عن الفترات الزمنية التي قضاها جالسا أو الفترات الزمنية التي تخلو من أفعال إيذاء الذات. وتتم مراقبة السلوك بصورة متواصلة خلال مرحلة المعالجة حتى يتبين أن للتدخل تأثيرا حيث يستقر معدل السلوك.

ويتحقق ضبط تجريبي أكثر في مرحلة A ثانية. ويكون ذلك في العادة، بانقطاع المعالجة، ولكن في بعض الحالات تعتبر مرحلة A الثانية قلب / عكس المعالجة بما يعزز سلوكاً يتعارض مع الاستجابة المرغوبة. إن إيقاف إعطاء القطع الرمزية للفتاة يعني إيقاف المعالجة، بيد أن إعطاءها القطع الرمزية عندما تغادر مقعدها هو قلب / عكس المعالجة. وفي كل حالة يتوقع أن تسبب إزالة المعالجة، عودة السلوك إلى المستوى الأصلي (خط القاعدة الأول). وثمة عيب في إنهاء التجربة مع المرحلة A الثانية، إذ يتم ترك الزبون في الحالة نفسها قبل بدء التجربة. ولأسباب أخلاقية، ولإضافة قوة إلى التصميم بتكرار الإجراء يتم إعادة مرحلة العلاج B مرة أخرى. وهناك أنواع مختلفة للتصميم (ABAB). فيمكن اختبار أكثر من معالجة – مثلاً استخدام صيغة (ABCACB) حيث A حالة اللا معالجة وتكون كل من B و C معالجات متناوية.

(*) يرجى العودة إلى مسرد شرح المصطلحات الواردة لاستيضاع مفهوم خط القاعدة (baseline) – (المراجع)

الشكل (9.5): (تصميم ABAB استخدام التعزيز الفارقي لتقليص وقت ارتداء الملابس صباحا لطفل في الثامنة من العمر)



ويوضح الشكل 9.5 تصميم ABAB. ففترات الصباح كانت صعبة في الغالب لكل أسرة طفل في الثامنة يعاني من إعاقة في تطوره وهو (كيرت)، إذ كان يستغرق ساعتين كي يرتدي ملابسه في الصباح. ولم تساعد وسائل التذكير المتواصل والتوسلات والتوبيخ العرضي. فمرة واحدة خلال أسبوع القياس في خط القاعدة A ارتدى كيرت ملابسه في (4) دقائق فقط، وهكذا أدرك والده أنه قادر على ارتداء ملابسه بنفسه في وقت معقول.

وخلال المرحلة (B) كانت توضع ملابس كيرت في الحمام كل صباح حيث يجرى إيقاظه كما يعمل منبه توقيت المطبخ لمدة عشر دقائق. ويحصل كيرت على ملصق أحمر ليضعه على اللائحة فوق بابه إذا ما ارتدى ملابسه قبل أن يدق جرس المنبه. لقد خوله الشريط مشاهدة التلفزيون ذلك المساء. وان لم ينته خلال (10) دقائق، فعليه البقاء في الحمام حتى يرتدي ملابسه ولا يسمح له بمشاهدة التلفزيون ذلك اليوم. لقد حظي كيرت بالإطراء عندما استطاع بلوغ ذلك المعيار، وأهمل عندما لم يستطع.

وأثناء أسبوع خط القاعدة اخذ كيرت معدل (59) دقيقة ليرتدي ملابسه. وخلال (17) يوما من المرحلة B الأولى، كان كيرت، على مستوى المعيار تسع مرات وكان المعدل (10) دقائق. وفي ثلاثة قياسات، تم تسجيل (صفر) دقيقة، لأنه استيقظ وارتدى ملابسه قبل أن يستيقظ والده.

خلال الأيام الستة التالية، وضعت ملابس كيرت خارج الحمام، مع حذف المؤقت

والسماح له بمشاهدة التلفزيون ودون اهتمام بزمن استغرقه في ارتداء ملابسه. كان وقت ارتداء ملابسه قد بلغ معدل (23) دقيقة خلال المرحلة A الثانية. وعندما أعيدت المعالجة مرة أخرى لمدة (12) من (13) يوماً كان وقت ارتداء كيرت لملابسه قد بلغ معدل (8) دقائق.

ونظراً لأنه كان يصحب المعالجة دائماً تغيير في زمن الارتداء، فقد أقيمت صلة موثوقة بين المعالجة وزمن الارتداء. إن العنصر الرئيسي في تصميم (ABAB) هو العودة إلى المستويات الأولى من السلوك لدى العودة إلى شرط خط القاعدة. والافتراض بان التدخل هو سبب التغيير يمكن أن يضعف، بشكل كبير، تحت بعض الظروف. فإذا كان هناك تباين كبير خلال فترة خط القاعدة، فبوسع المرء أن يشير إلى أن سلوك المعالجة البعدية لم يكن مختلفا عن سلوك المعالجة القبلية. وإذا كان التحول من المعالجة (B) إلى خط القاعدة (A) غير واضح فإن هذا يضعف الكثير من قوة تصميم (ABAB). ولن يبرز في خط القاعدة الثاني، غالباً، تطرف السلوك الأول نفسه، وليس هناك في بعض الحالات عودة إلى المعدلات السابقة. وفي بعض الحالات، قد تسبب عوامل أخرى، مثل التاريخ والنضوج، التأثير الملاحظ.

لقد أشار (Kazdin, 1980) إلى أن العودة إلى خط القاعدة، قد يكون غير أخلاقي تحت بعض الظروف. فهل تريد حقا من الطفل المتوحّد، العودة إلى المستويات السابقة لضرب الرأس فقط لمجرد إعطاء الدليل بان المعالجة هي التي سببت التقليص؟ وبسبب هذه العوامل المحتملة مع تصميم (ABAB)، تم تطوير تصميم خط القاعدة المتعدد. ولهذا التصميم فائدة سريرية في استمرار أي تحسن دون العودة إلى سلوك مرغوب أقل.

تصميمات خط القاعدة المتعدد Multiple - Baseline Designs

يتم في تصميم خط القاعدة المتعدد، قياس أنواع السلوك المختلفة في الوقت ذاته خلال ظرف خط القاعدة. فمثلاً، قد يسجل الباحث عدد المرات التي تحدث فيها تلميذ في الصف دون استئذان، وعدد أوراق الواجبات التي أتمها، وعدد المرات التي ضرب فيها التلميذ طفلا آخر. ويمكن تسجيل ذات السلوك لعدة زبائن، مثل مقدار الاتصال بالعين(*) مع المدرس، بالنسبة لطفلين أو أكثر في صف ما قبل المدرسة، خلال مرحلة خط القاعدة. وفي كلا الحالتين تظل المعالجة ذاتها (B) قائمة عبر كل الظروف أو الطلبة. وينشأ الضبط التجريبي في خط القاعدة المتعدد من بدء المعالجة، في نقطة زمنية مختلفة، لكل سلوك و / أو لكل شخص معني، وليس من العودة إلى خط القاعدة. وهكذا، فبعد تحديد خط القاعدة، يتم البدء بمعالجة السلوك (1) ويتم الاستمرار بخط القاعدة لكل من السلوك (2) و (3). وعندما يتم البدء بمعالجة السلوك (2) يتم الاستمرار بمعالجة السلوك (1) وخط القاعدة للسلوك (3). وأخيراً يتم البدء بمعالجة السلوك (3). ومن المتوقع أن

(*) الاتصال بالعين (eye contact): مصطلح يشير إلى تصويب النظر نحو من ينظر إليك- (المراجع)

يتغير كل سلوك في الاتجاه المرغوب في النقطة التي تبدأ فيها المعالجة، وليس قبلها أو بعدها.

وهكذا يستخدم تصميم خط القاعدة المتعدد AB كوحدة أساسية له. فإذا كانت هناك حادثة، خلافاً للمعالجة، هي السبب الفعلي للتغييرات فإنها لابد أن تؤثر على كل الأطفال أو على كل أنواع السلوك في ذات النقطة الزمنية. إن أحد افتراضات هذا التصميم، هو أن المعالجة تؤثر على أنواع السلوك المختلفة بصورة محددة. ولا يتوقع من تعزيز سلوك واحد (إكمال مسائل الحساب) أن يزيد استجابة أخرى (معدل القراءة). ويجب أن تكون أنواع السلوك، أو المواقف، مستقلة (غير مترابطة) كي تبين دراسة خط القاعدة المتعدد تأثيرات قابلة للتفسير. وفي الواقع، قد يكون من الصعب الحصول على استقلال السلوك. فتعديل سلوك واحد (مثل الحديث في الصف) قد يؤثر على أنواع السلوك المستهدفة الأخرى (إكمال الواجبات في الوقت المحدد). وينصح القارئ بالرجوع إلى كتاب (Kazdin & Kopel, 1975) لإطلاع أوسع على مشكلات خط القاعدة المتعدد.

مقارنة تصميمي الفرد الواحد والمجموعة

Comparison of Single-Subject and Group Designs

يهدف الباحث في كل من تجارب الفرد الواحد والمجموعة لحو إقامة صلة، لا لبس فيها، قدر الإمكان، بين تفعيل المتغير المستقل (المعالجة) وأثره على المتغير التابع (السلوك). ففي تصميمات المجموعة، يستبعد التخصيص العشوائي للأفراد في المجموعة التجريبية أو المجموعة الضابطة، العديد من التفسيرات المنافسة للفروقات الملاحظة بعد المعالجة. ويمكن تقييم تأثيرات المعالجة (بين المجموعات) ذات الصلة بتأثيرات التغير (داخل المجموعات) عن طريق استخدام الاختبارات الإحصائية المناسبة. وتحدد هذه الاختبارات ما إذا كانت الصدفة وحدها تفسيراً موثوقاً للنتائج. ويستخدم تصميم الفرد الواحد طرقاً أخرى لبناء الثقة. فالباحث يضبط مقدار الوقت الذي تكون فيه مرحلتا خط القاعدة المعالجة فاعلتين، ومقدار تمديد فترة خط القاعدة حتى يستقر السلوك. ولأجل الحصول على تفسير غير ملتبس لابد أن يكون خط القاعدة محدداً نسبياً أو يكون الميل في الاتجاه المعاكس مما هو متوقع بعد المعالجة. وتكمن إحدى عيوب ضبط طول مرحلة المعالجة لدى الباحث، في النـــزوع نحو استمرار المعالجة حتى "يحدث شي ما". فإذا لم يتبع بداية المعالجة تغير في السلوك فيمكن أن يكون هناك متغير غير تجريبي آخر، تسبب في التغير الملاحظ.

تتجنب التصميمات التجريبية لفرد واحد مصدر خطأ في تصميمات المجموعة – أي التغير داخل الأفراد. فكل فرد يخدم كضابط لنفسه، لذا لا تكون المقارنة مشكلة. فوسيلة الضبط الرئيسية هي الإعادة، وهو جانب قلما يندمج في تصميمات المجموعة. وينطوي تصميم ABAB على إعادة واحدة باستخدام الفرد ذاته، بينما يعيد تصميم خط القاعدة المتعدد أكثر من معالجة. فإعادة تصميم

خط القاعدة المتعدد يجعل الأمر أقل احتمالاً، من أن تُنسب التأثيرات التي تعزى إلى المعالجة، إلى حادثة دخيلة أو متغيرات الأفراد (Birnbauer, Peterson, & Solnick, 1974).

بوسع البحث الخاص بالفرد الواحد، المصمم جيداً، أن يلبي معايير الصدق الداخلي. وعلى أية حال، فإن مسألة الصدق الخارجي - أي تعميم الاستنتاجات التجريبية - لا يمكن الإجابة عليها بسهولة عن طريق تصميمات تستخدم فقط فرداً، أو أفراداً قليلين. وبوسع المرء أن يبين أن السماح لمراهق مضطرب سلوكيا أن ينصت إلى موسيقى الروك لدى إكماله واجباته، سوف يزيد من مقدار العمل المدرسي الذي يتمه ذلك المراهق، إلا أنه كيف يمكننا أن نحدد ما إذا كانت المعالجة ناجحة مع مراهقين آخرين أم مع مراهقين مضطربين سلوكيا؟ ومع أن أية دراسة خاصة لفرد واحد، ستكون متدنية في صدقها الخارجي، إلا أن عدد الدراسات المشابهة التي تصف الأفراد والظروف والمعالجات، ستبني الحالة من أجل تطبيق واسع لتأثيرات المعالجة الخاصة. ولأجل مناقشة اكمل لتصميمات الفرد الواحد، انظر (Kazdin, 1982).

الخلاصة SUMMARY

التجربة هي الشكل الأكثر دقة ورغبة في البحث العلمي. والظروف الضابطة التي تميز التجربة، تتيح تحديد العلاقات الوظيفية المؤكدة بين الظواهر المعنية لدى التربويين. ويتمتع الباحثون الذين يضبطون الظروف التي تحدث فيها حادثة معينة بمزايا بارزة على المراقبين الذين يراقبون أو يدرسون، ببساطة، الحادثة تلك دون ضبط:

1- بإمكانهم تفعيل أو تغيير الظروف بشكل منتظم وملاحظتهم للاختلافات في النتائج.

2- يمكنهم جعل الحادثة أن تقع عندما يكونون على استعداد لإجراء ملاحظات وقياسات دقيقة.

3- يمكنهم تكرار ملاحظاتهم تحت الظروف ذاتها، من أجل تأكيدها، ويمكنهم وصف هذه الظروف بحيث يتسنى للباحثين إعادتها وإجراء تدقيق مستقل للنتائج.

ينبغي أن يكون للتصميم البحثي صدق خارجي وداخلي. ويهتم الصدق الداخلي بالسؤال التالي: هل سببت المعالجة التجريبية (المتغير المستقل) التغييرات المرصودة في المتغير التابع، أو هل كانت هذه التغيرات بسبب بعض المتغيرات غير المعنية؟ ولأي تصميم بحثي صدق خارجي إذا كان بالإمكان تعميم نتائج البحث في أوضاع أخرى، وتعريفات إجرائية، ومجتمعات إحصائية.

توفر التصميمات التجريبية الحقيقية أفضل ضبط للصدق الداخلي. وفي هذه التصميمات يتم تخصيص الأفراد عشوائيا في المعالجات. ومع التصميمات شبه التجريبية، يمكن للباحث ضبط

المعالجة غير أنه لا يسعه تخصيص الأفراد للمعالجات عشوائيا. لقد أثبتت التصميمات شبه التجريبية، التي تدرس اثر المعالجة على فرد واحد، فائدتها في البحوث السلوكية.

مفاهيم أساسية Key Concepts

ABAB design	تصميم ABAB
analysis of covariance (ANCOVA)	تحليل التباين الاقتراني (ANCOVA)
comparison group	مجموعة المقارنة
confounding variable	متغير مربك / مشوش
control group	مجموعة ضابطة
control group time series design	تصميم السلسلة الزمنية ذو المجموعة الضابطة
control of variables	ضبط المتغيرات
controlling situational differences	ضبط الفروقات الموضعية / الموقفية
counterbalanced design	تصميم الموازنة / التدوير
differential selection	اختيار فارقي / تمييزي
ecological validity	صدق بيئي
experimental design	تصميم تجريبي
experimental group	مجموعة تجريبية
experimental research	بحث تجريبي
experimenter bias	تحيز تجريبي
external validity	صدق خارجي
external validity of operations	الصدق الخارجي للعمليات
extraneous variable	متغير دخيل
factorial design	تصميم عاملي
Hawthorne effect	أثر / ظاهرة هوثورن
history (as internal-validity problem)	التاريخ (كمشكلة صدق داخلي)
homogeneous selection	اختيار متجانس
implementation threat	تهديد / مخاطر التنفيذ
interaction in factorial design	التفاعل في التصميم العاملي
interaction of subject characteristics and treatment	تفاعل سمات الأفراد والمعالجة
internal validity	صدق داخلي
John Henry effect	أثر / ظاهرة جون هنري
law of the single significant variable	قانون المتغير الواحد المتميز

manipulation of independent variable	تفعيل المتغير المستقل
maturation (as internal-validity problem)	النضج (كمشكلة صدق داخلي)
measuring instruments (as internal-validity problem)	أدوات قياس (كمشكلة صدق داخلي)
multiple-baseline design	تصميم خط القاعدة المتعدد
nonrandomized control group, pretest-posttest design	تصميم المجموعة الضابطة غير العشوائية ذو الاختبارين القبلي والبعدي
one-group pretest-posttest design	تصميم المجموعة الواحدة ذو الاختبارين القبلي والبعدي
one-group time series design	تصميم السلسلة الزمنية ذو المجموعة الواحدة
population validity	صدق المجتمع الإحصائي
preexperimental research	البحث قبل التجريبي
pretest sensitization	حساسية الاختبار القبلي
pretesting (as internal-validity problem)	إجراء الاختبار القبلي (كمشكلة صدق داخلي)
quasi-experimental research	البحث شبه التجريبي
random assignment (randomization)	التخصيص العشوائي (العشوائية)
randomized matched subjects, posttest-only control group design	تصميم المجموعة الضابطة حسب تناظر عشوائي للأفراد، ذو اختبار بعدي فقط
randomized matching	التناظر العشوائي
randomized subjects, posttest- only control group design	تصميم المجموعة الضابطة حسب تخصيص عشوائي، ذو اختبار بعدي فقط
randomized subjects, pretest-post test control group design	تصميم المجموعة الضابطة حسب تخصيص عشوائي ذو اختبارين قبلي وبعدي
random selection	اختيار عشوائي
single-subject research	البحث ذو الفرد الواحد
Solomon four-group design	تصميم سولومون ذو المجموعات الأربع
Solomon three-group design	تصميم سولومون ذو المجموعات الثلاث
static group comparison	مقارنة المجموعة الثابتة
statistical equivalence	التكافؤ الإحصائي
statistical regression (as internal- validity problem)	الارتداد الإحصائي (كمشكلة صدق داخلي)
true experimental research	البحث التجريبي الحقيقي
using subjects as their own controls	استخدام الأفراد كضابطين لأنفسهم

تمارين EXERCISES

1- من مجموعة طلبة مسجلين في الدراسات اجتماعية في مدرسة ثانوية، اختار باحث بصورة عشوائية 60 طالبا. ثم قسم الطلبة إلى مجموعتين بتخصيص عشوائي هو 30 طالباً للمجموعة أ يتبعون منهج الدراسات الاجتماعية التقليدية، و 30 طالبا للمجموعة ب يتبعون البرنامج الجديد المصمم للتعامل مع التاريخ لدى بعض المجموعات العرقية. وقورنت المجموعتين في نهاية الفصل الدراسي حسب مقياس مصمم لقياس الاتجاهات نحو المجموعات العرقية. في هذه الدراسة، حدد ما يلي:

أ- المتغير المستقل.

ب- المتغير التابع.

ج- المجموعة الضابطة.

د- المجموعة التجريبية

هـ- الطرق المستخدمة لضبط الفروقات بين المجموعتين.

و- التصميم البحثي المستخدم.

2- ادرس السؤال البحثي التالي: هل يغير تدريس اللغة الفرنسية للصف الأول من خلال الطريقة الشفوية - السمعية، بدلا من الطريقة النحوية - التحويلية، من أداء التلميذ في اختبار نهاية السنة المقنن في النحو والقراءة والمفردات؟

أ - صمم التجربة المثالية للإجابة على هذا السؤال، بافتراض عدم وجود قيود إدارية أو أية قيود أخرى.

ب - صمم التجربة الأكثر احتمالاً لحاجتها في الوضع النموذجي لمدرسة ثانوية.

ج - وضح الفوائد النسبية للتصميم التجريبي المثالي (التمرين 2 أ) مقارنة بالتصميم في التمرين 2 ب.

3- ما هو الفرق بين الاختيار العشوائي والتخصيص العشوائي؟ وكيف يرتبطان بالصدق الداخلي والصدق الخارجي؟

4- قيّم التصميمات البحثية التالية فيما يخص الطرق المستخدمة والضبط المتوفر. اعط اقتراحات للوصول إلى تحسينات عند الحاجة.

أ - أراد باحث التأكد ما إذا كان تجميع متجانس يحسن من التعلم في أول مقرر في علم الأحياء. وخصص الباحث واحدة من مدرستين ثانويتين في مدينة صغيرة كي تكون مدرسة تجريبية والأخرى ضابطة. وكان لكلا المدرستين العدد نفسه من الطلبة في كل من أربعة صفوف علمية. وفي المدرسة التجريبية، تم جمع الطلبة بصورة متجانسة على أساس معامل الذكاء ودرجات اختبارات التحصيل في العلوم. وفي

المدرسة الضابطة، تم وضع الطلبة في صفوف بشكل عشوائي. وفي نهاية السنة أعطى جميع الطلبة اختباراً مقنناً في علم الأحياء. لقد أوضحت الاختبارات الإحصائية تفوق المجموعة التجريبية في الاختبار. استنتج الباحث أن التجميع المتجانس أدى إلى تعلم أعلى في علم الأحياء.

ب- كانت مدرسة التاريخ قلقة بسبب افتقار طلابها إلى معرفة دولتهم وحكوماتهم الوطنية والأحداث الجارية. فقررت أن تجرب طرقاً ومواد جديدة كي ترى ما إذا كان بوسعها الحصول على تحسن معين. في الصفين أ و ب، أدخلت مواد جديدة. وفي الصفين ج و د، استخدمت الطرق التقليدية. وأجرى للصفين أ و ب اختبار قبلي واختبار بعدي، بينما أعطى للصفين ج و د اختبار بعدي فقط. ولدى إجراء المقارنات في الاختبار البعدي وجدت أن الصفين أ و ب أكثر تفوقا. وقد عزي أداؤهم المتفوق إلى المواد والطرق الجديدة.

5- صمم التجربة المثالية لاختبار الفرضية التالية: إذ شاهد أطفال أفلاماً ذات تفاعل عرقي منسجم، فإنهم سيبدون اتجاهات أكثر إيجابية نحو الأقليات العرقية من الأطفال الذين يشاهدون أفلاماً تشير إلى الصراع العرقي.

6- بالعودة إلى المسألة البحثية في التمرين 2، افترض أنك تريد أيضاً أن تعرف ما إذا كانت لطريقتي تدريس الفرنسية تأثيرات فارقية على الأولاد والبنات؟ حدد التصميم التجريبي الذي يسمح لك بالإجابة على هذا السؤال في الوقت ذاته.

7- افترض أن باحثا كان قد استخدم طريقتين للتدريس ($أ_2$، $أ_1$) مع مجموعتين من الطلبة ($ب_2$ ، $ب_1$) وبمستويات مختلفة من دافع التحصيل. وقورنت المجموعتان في اختبار تحصيل في نهاية الدراسة. الأوساط معروضة أدناه. ما التفسير الذي ستقدمه عن هذه النتائج؟

	$أ_1$	$أ_2$
$ب_1$	35	15
$ب_2$	15	35

8- ماذا يجب أن يحدث لإقامة موثوقية النتائج في تصميم الفرد الواحد؟

9- حدد التهديدات / المخاطر للصدق الداخلي في كل من الدراسات التجريبية التالية:

أ - في بداية السنة الدراسية تقيم مدرّسة ابتدائية مواقع قراءة في غرفتها مع مواد قراءة إضافية لكل موقع. وفي نهاية السنة ذكرت أن اهتمام صفها في القراءة كان أعلى مما كان في صفوفها السابقة. لقد أوصت بأن يقيم جميع المدرسين مواقع قراءة.

ب- أراد مدرس بحث فاعلية مواد تدريسية جديدة مصممة لزيادة المهارات اللفظية للطلبة. واستخدم واحداً من صفين، المواد الجديدة، والآخر كتيبات العمل(*) التقليدية. وأعطي لكلا الصفين اختبار بالمهارات اللفظية في نهاية الفصل.

الأجوبة ANSWERS

1- أ- نوع من منهج الدراسات الاجتماعية
ب- درجات على مقياس الاتجاهات العرقي
ج- المجموعة أ، المنهج الحالي
د- المجموعة ب، المنهج مع التاريخ العرقي
هـ- اختيار عشوائي للعينة من مجتمع إحصائي، وتخصيص عشوائي للعينة لكل من المجموعة التجريبية والضابطة.
و- التصميم 3، تصميم المجموعة الضابطة حسب تخصيص عشوائي، ذو اختبار بعدي فقط

2- أ- استخدم التصميم 3 - أي تخصيص طلبة الفرنسي عشوائيا للمجموعة النحوية - التحويلية (الضابطة) أو الشفوية - السمعية (التجريبية). حافظ على ذات الظروف، والوقت المطلوب، والمدرسين، والتسهيلات الصفية لكلا المجموعتين بحيث تكون طريقة التدريس هي المختلفة فقط. اجر اختباراً في نهاية السنة وقارن تحصيل المجموعتين.

ب- خصص عشوائيا صفوفاً تامة لطلبة للصف الأول - فرنسي لطريقتي التدريس. لكل مدرس عدد متساو من نوعي الصفوف.

ج- في التصميم المثالي، يجرى التحكم بمخاطر الصدق الداخلي والخارجي، بشكل أفضل، من خلال عشوائية توزيع الطلبة. التصميم في التمرين 2 ب قد ينطوي على مشكلات عدم التكافؤ بين الأفراد قبل إعطاء المعالجة، حيث فروقات درجات الاختبار يمكن أن تعزى إلى عوامل غير الاختلاف في المعالجة.

3- يستخدم الاختيار العشوائي إجراء الصدفة لسحب عينة من مجتمع إحصائي. ولأنه يتناول مسألة مدى تعميم النتائج المستمدة من العينة على المجتمع الإحصائي الذي سحبت منه

(*) كتيبات العمل (workbook): كتيب يشتمل على معلومات موجزة للحقل بالإضافة إلى مجموعة من الأسئلة والتمارين مع تعليمات للتنفيذ - (المراجع)

العينة فهو يعد استراتيجية لزيادة الصدق الخارجي. أما التخصيص العشوائي، فيستخدم إجراء الصدفة لتخصيص الطلبة المتيسرين لتجربة معينة للمعالجة. إنها استراتيجية لزيادة الصدق الداخلي.

4- أ- نظراً لأن الباحث لا يستطيع تخصيص الطلبة عشوائيا للمدارس الثانوية، فإن هناك مخاطر عدة للصدق الداخلي. فالطلبة في المدرسة التجريبية قد يكونون أذكى، أو لديهم خلفية أكثر في العلوم من الطلبة في المدرسة الضابطة. إن الفروقات في نوع تدريس علم الأحياء في المدارس لم تخضع للضبط. ونظراً لأن الباحث استخدم فقط المدارس في مدينة صغيرة، فإن نتائج الدراسة لا يمكن تعميمها على المدارس الثانوية في أوضاع مختلفة. ولتحسين الدراسة يمكن للباحث مقارنة تحصيل العلوم في البدء ودرجات معامل الذكاء للمدارس كي يرى ما إذا كانت المجموعات متكافئة قبل المعالجة. وباستخدام عدة مدارس ثانوية مع صفوف في كل مدرسة خصصت بشكل عشوائي لظروف تجريبية، سيكون ضبط العوامل الخاصة بمدرسة معينة.

ب- الصفوف التي لم تخصص عشوائيا قد لا تكون متكافئة. ويمكن إجراء اختبار قبلي لتحديد التكافؤ، لكنه أعطى فقط للمجموعات التجريبية. إن إجراء الاختبار القبلي للمجموعات التجريبية، قد يثير حساسية المجموعات ويؤثر على الفروقات المكتشفة. ينبغي أن تخصص الصفوف عشوائياً للمعالجات حتى وإن لم يتوفر تخصيص الطلبة عشوائياً.

5- في التجربة المثالية، سيخصص الطلبة عشوائياً للمجموعتين. وسوف تستخدم نتائج قياس اتجاهات المعالجة البعدية نحو الأقليات العرقية، لمقارنة المجموعتين التجريبية والضابطة.

6- يحتاج هذا السؤال إلى تصميم عاملي مع تخصيص نصف الأولاد ونصف البنات عشوائيا للشروط الضابطة والتجريبية.

7- يبدو أن هناك تفاعلاً بين دافع التحصيل ونوع التدريس. فالطلبة مع دافع التحصيل في المستوى ب$_1$ كانوا أفضل مع الطريقة أ$_1$، بينما الطلبة في المستوى ب$_2$ أدوا بشكل افضل مع الطريقة أ$_2$. يمكن اختبار دلالة التفاعل باستخدام اختبار F. وليس هناك تأثير كلي للدافع أو طريقة التدريس لأن أوساط أ$_1$ و أ$_2$ و ب$_1$ و ب$_2$ هي نفسها.

8- لابد من وجود تغيير واضح / غير ملتبس في السلوك متى ما كان هناك تغيير في المعالجة.

9- اختيار فارقي، تنفيذ، تاريخ معاصر، اتجاهات الأفراد.

REFERENCES المصادر

Birnbauer, L., Peterson, C., and Solnick, J. (1974). Design and interpretation of studies of single subjects. *American Journal of Mental Deficiency*, 79, 191-203.

Campbell, D. T., and Stanley, J.C. (1966). *Experimental and quasi- experimental designs for research.* Boston: Houghton Mifflin.

Heward, W., Dardig, J.C., and Rossett, A. (1979). *Working with parents of handicapped children.* Columbus, OH: Charles E. Merrill.

Kazdin, A. (1980). *Research design in clinical psychology.* New York: Harper & Row.

Kazdin A.E. (1982). *Single case research designs: Methods for clinical and applied settings.* New York: Oxford University Press.

Kazdin, A., and Kopel, S. (1975). On resolving ambiguities of the multiple baseline design: Problems and recommendations. *Behavior Therapy*, 6, 601-608.

Kempthome, O. (1961). The design and analysis of experiments with some reference to educational research. In R.O. Collier and S.M. Elam (Eds.), *Research design and analysis: Second annual Phi Delta Kappa symposium on Educational research.* Bloomington, IN: Phi Delta Kappa.

Mill, J.S. (1846). *A system of logic.* New York: Harper & Brothers.

Rice, J.M. (1897). The futility of the spelling grind. *Forum*, 23, 163-172, 409-419.

Rosenthal, R. (1985). From unconscious experimenter bias to teacher expectancy effects. In J.B. Dusek (Ed.), *Teacher expectancies* (pp. 37-65). Hillsdale, NJ: Lawrence Erlbaum.

Schoen, H.L. (1976). Self-paced mathematics instruction: How effective has it been? *Arithmetic Teacher;* 23, 90-96.

Smith, M.L., and Glass, G.V. (1987). *Research and evaluation in education and the social sciences.* Englewood Cliffs, NJ: Prentice Hall.

Solomon, R.L. (1949). On extension of control group design. *Psychological Bulletin*, 46, 137-150.

Thorndike, E.L. (1924). Mental discipline in high school subjects. *Journal of Educational Psychology*, 15; 1-22, 83-98.



البحث العلّي (*) - المقارن

Causal - Comparative Research

INSTRUCTIONAL OBJECTIVES أهداف تعليمية

بعد دراسة هذا الفصل سيكون بوسع الطالب أن:

1- يصف البحث العلّي – المقارن ويقارنه بالبحث التجريبي.

2- يذكر الشروط المطلوبة لاستنتاج العلاقة العلّية.

3- يصف التفسيرات البديلة في البحث العلّي – المقارن ويحدد الحالات التي تبدو معقولة أم لا.

4- يصف طرق الضبط الجزئي ويحدد الحالات التي تكون فيها مفيدة.

5- يصمم الاستقصاء / البحث العلّي – المقارن.

6- يحدد الأسئلة التي يكون فيها البحث العلّي – المقارن طريقة الاختيار.

(*) ثمة تمييز بين العلّة (Cause) والسبب (Reason). فالعلّة هي ما يُحدث التأثير أو النتيجة، وإذا كانت وحدها أو واحدة في سلسلة، فلابد من وجودها لحدوث التأثير المنطقي.
أما السبب (Reason) فإنه ما يشرح وقوع وطبيعة التأثير بدلالة التفكير الإنساني، وليس بالاعتماد على العوامل الموضوعية أو الخارجية. (المراجع)

عندما نستقصي سؤالاً تربوياً مثل، لماذا يعتبر بعض الأطفال قارئين أفضل من أطفال آخرين؟ ولماذا يصبح بعض الشباب جانحين بينما لا يصبح غيرهم كذلك؟ نجد أن بعضاً من أسئلتنا يمكن بحثها عن طريق البحث التجريبي، بينما لا يمكن بحث أسئلة أخرى على هذه الشاكلة. وإذا أردنا بحث تأثير متغيرات مثل البيئة البيتية، والدافع، والذكاء، وعادات القراءة لدى الأبوين، والعرقية، وما إلى ذلك، فيتعذر علينا أن نخصص الطلبة عشوائيا في أصناف مختلفة لهذه المتغيرات. فالمتغيرات المستقلة كهذه تدعى "المتغيرات المنسوبة". والمتغير المنسوب (attribute variable) هو خاصة / سمة يمتلكها فرد قبل بدء الدراسة.

وخلافاً لذلك، فإن المتغير المستقل الذي يستطيع الباحث تفعيله هو "متغير فعال / نشيط"(active variable). وبوسع الباحث أن يحدد الطلبة الذين ستكون لديهم حرية الوصول إلى مختبر الحاسوب، والذين لا يستطيعون ذلك، والذين سيستخدمون البرنامج (أ) لدراسة وحدة في مادة الجبر، والذين سيستخدمون البرنامج (ب). وعندما ينطوي الأمر على متغيرات مستقلة فعالة فبوسعنا استخدام البحث التجريبي أو شبه التجريبي. أما عندما يكون لدينا متغيرات مستقلة منسوبة فعلينا أن نتحول إلى البحث العلّي- المقارن (ويدعى أحيانا البحث اللاحق / الراجع للحادث ex post facto research).

جاءت التسمية (ex post facto) من اللاتينية التي تعني (من بعد الحقيقة) وتفيد في إيضاح أن البحث المعني يجرى بعد حدوث التغيرات في المتغير المستقل في السياق الطبيعي للأحداث. فالباحث لا يملك التحكم المباشر بالمتغيرات المستقلة، وذلك لأنه لا يمكن تفعيلها بسبب طبيعتها أو لأنها وقعت مسبقاً. ونظراً لأن هذا البحث، غالباً، ما يجرى لإيجاد علل الأشياء التي حدثت عن طريق مقارنة الظروف القائمة سابقاً، فإنه أصبح يدعى البحث العلّي- المقارن.

أما النموذجان الأساسيان للبحث العلّي- المقارن فهما: 1) البدء مع أفراد يختلفون حسب متغير مستقل واختيار الفرضيات الخاصة بكيفية اختلافهم في متغيرات تابعة، 2) البدء مع أفراد يختلفون في متغير تابع واختبار فرضيات تخص متغيرات مستقلة محتملة. ومثال عن النوع الأول يكون باختبار الفرضية القائلة بان التهرب من المدرسة (متغير تابع) يلاحظ بشكل أكبر في العائلات واحدة الأبوين من العائلات ذات الأبوين (متغير مستقل). ومثال عن النوع الأخر، يختبر الفرضية القائلة أن الذين يتخرجون من الثانوية يختلفون عن الذين يرسبون (متغير تابع) في الدافعية، ووضوح الأهداف، والانضباط الذاتي (متغيرات مستقلة).

ويحوز الباحثون على التباين الذي يحتاجونه، ليس عن طريق التفعيل المباشر للمتغير ذاته، بل عن طريق اختيار الأفراد الذين يكون فيهم المتغير حاضراً أو غائباً، قوياً أو ضعيفاً وما إلى ذلك. انهم يقدمون للأطفال ذوي التلف الدماغي والذين لا يعانون فيه الواجب الإدراكي نفسه، أو يقارنون أداء الأطفال ذوي معامل الذكاء العالي والمنخفض حسب مقياس القلق ذاته.

مقارنة أسلوبي العلّي – المقارن والتجريبي
CAUSAL - COMPARATIVE AND EXPERIMENTAL APPROACHES COMPARED

يتركز الاهتمام في كلا نوعي البحث على اكتشاف أو إقامة علاقات بين المتغيرات في بيانات الباحث. ويمكن للبحث العلّي – المقارن والبحث التجريبي اختبار فرضيات تخص العلاقة بين متغير مستقل (X) ومتغير تابع (Y). ففي المنطق الأساسي، تتشابه الأساليب التجريبية والعلّية – المقارنة. وكذلك، فإن الكثير من المعلومات من النوع ذاته التي توفرها التجربة يمكن توفيره من خلال التحليل العلّي – المقارن.

ويمكن، على أية حال، الحصول مع التجربة على أدلة أكثر إقناعا للعلاقات العلّية أو الوظيفية بين المتغيرات، مما يمكن الحصول عليه مع الدراسات العلّية – المقارنة. ويتم ضبط تأثيرات المتغيرات الدخيلة في تجربة معينة عن طريق الظروف التجريبية، والمتغير المستقل الذي يفترض أن يكون سابقاً، يجرى تفعيله بصورة مباشرة بغية تأكيد تأثيره على المتغير التابع. فإذا لوحظ أن (Y) تتغير بالتلازم مع تغير (X) في هذا الوضع الخاضع للضبط، فإنه يتوفر للمرء دليل على صحة علاقة (السابق – التالي) المفترضة بين (X) و (Y). أما في البحث العلّي– المقارن، من ناحية أخرى، فإن الباحث لا يستطيع ضبط المتغيرات المستقلة بالتفعيل أو بالعشوائية. فالتغيرات في المتغير المستقل أو المتغيرات قد حدثت قبلاً. ويواجه الباحث بمشكلة محاولة تحديد المتغيرات السابقة للنتيجة الملاحظة. وبسبب الافتقار إلى الضبط، فإن ما هو أكثر خطورة، الاستنتاج بأن هناك علاقة أصيلة بين (X) و (Y) في دراسة علّية – مقارنة.

دعنا نوضح الفرق بين أسلوب العلّية – المقارنة والأسلوب التجريبي عن طريق تفحص هذين الأسلوبين إزاء مسألة بحثية واحدة. خذ مسألة تأثير قلق الطلبة في وضع اختبار تحصيل على أدائهم. فأسلوب العلّية – المقارنة سوف ينطوي على قياس مستوى القلق الموجود مسبقاً في وقت الاختبار، ثم مقارنة الأداء لدى الطلبة ذوي القلق العالي والطلبة ذوي القلق المنخفض. ويكمن ضعف مثل هذا الأسلوب في كون أن المرء لا يمكنه بالضرورة أن يستنتج بأن قلق الطلبة هو الذي ولد الفرق الملاحظ في أداء اختبار التحصيل. فربما تأثرت كلا مجموعتي الدرجات بعامل ثالث، مثل معرفة الموضوع الذي تم اختباره أو الذكاء العام. فالمعرفة أو الذكاء قد يكونان العلة الرئيسية لكل من مستوى القلق ونتائج اختبار التحصيل.

أما الأسلوب التجريبي للمشكلة ذاتها فينطوي على إجراء اختبار تحت ظرفين يتطابقان في كل جانب، عدا أن أحدهما يثير القلق والأخر محايد. وبوسع الباحث إثارة القلق بالقول للأفراد بأن تقديرهم النهائي يعتمد على أدائهم، أو أن الاختبار صعب جداً، أو أن الاختبار سوف يستخدم لتحديد غير الكفء. وسوف يقال للمجموعة المحايدة أن مجرد تعاونها مطلوب

للتجربة. وبوسع الباحث أن يخصص الأفراد عشوائياً لكلا الظرفين. فإن أدت المجموعة القلقة بشكل افضل من المجموعة المحايدة، يمكن الاستنتاج أن للقلق تأثيراً ميسّراً على أداء الاختبار. أن مثل هذا الاستنتاج يمكن أن يستمد بصورة مشروعة، بسبب الضبط الآتي من التخصص العشوائي للمجموعتين في المعالجات وبالتفعيل المباشر للمتغير المستقل من الباحث. فالقلق هو أحد المتغيرات القليلة التي يمكن أن تكون متغيراً مستقلاً فعالاً أو منسوباً. ويمكن للمرء أن يفعّله بصورة فعالة كما تم وصفه (الأسلوب التجريبي) أو يمكن للمرء أن يصنف الأفراد على أساس درجاتهم في مقياس القلق (الأسلوب العلّي - المقارن).

ويمكن، بطريقة أخرى، النظر إلى الدراسة العلّية - المقارنة كأسلوب عكسي للتجريب. فبدلاً من اخذ مجموعات متكافئة وتعريضها لمعالجات مختلفة، تبدأ الدراسة العلّية - المقارنة بمجموعات مختلفة سلفاً وتحاول تحديد العواقب أو السابقات لهذه الفروقات. إن مثل هذا الإجراء لا يوفر الضمانة، وهي أمر نموذجي في التجريب، التي تعد ضرورية لإقامة استنتاجات قوية حول العلاقات العلّية. فالباحث الذي يجد علاقة بين المتغيرات في دراسة علّية - مقارنة قد ضمن الدليل، فقط، لبعض التغير المصاحب / الملازم. ونظراً لأن الباحث لم يضبط (X) أو أية متغيرات محتملة أخرى، قد تحدد (Y) فإن هناك أساساً أقل لاستنتاج علاقة علّية بين (X) و (Y). فلكي تكون قادرا على استنتاج نوع علاقة من النمط (متقدم - نتيجة) بين المتغيرات، يجب أن تجمع الأدلة كي تبين أن (Y) لا يَسبق (X)، أو أن التأثير في (Y) لم يأت من عامل آخر على صلة بالعامل السابق المفترض. فمثلاً، إذا أعطينا لأطفال يعانون من تلف في الدماغ وأطفال دون تلف، اختبارا إدراكياً حسياً، فإن الفروقات في الأداء قد تعكس تأثيرات تلف الدماغ، أو إنها قد تعكس عوامل أخري كالفروقات في القلق المرتبط بأنواع معينة من المرض. أو دعنا نأخذ دراسة مبكرة قام بها (Rogerson & Rogerson, 1939) حيث ذكرت نتيجة مفادها أن مجموعة الأطفال الذين رضعوا رضاعة طبيعية خلال الطفولة قد أبدوا مستوى أداء في المدرسة الابتدائية أعلى مما فعلت مجموعة الأطفال الذين تغذوا من حليب الزجاجة. ولا يمكن الاستنتاج من نتيجة كهذه بأن الأداء في المدرسة يتحسن بالضرورة بالرضاعة من صدر الأم خلال فترة الطفولة. ومن المحتمل أن تكون العلاقة التي لاحظها الباحثان كانت نتيجة التغيرات في متغير واحد أو أكثر بما أثر على كل من نوع التغذية التي حصل عليها الأطفال الذين خضعوا للدراسة ومستوى الأداء الذي حققوه لاحقا في المدرسة. وذكر الباحثان أن الدراسة قد أجريت في عيادة تم تشجيع الرضاعة الطبيعية فيها، وأن الإخفاق في القيام بذلك كان نتيجة الصحة السيئة للطفل أو ألام أو كليهما. وهكذا، فالتفسير المعقول هو أن كلا من نوع الرضاعة التي حصل عليها الأطفال الصغار وأداءهم التالي في المدرسة قد تأثرت بالصحة - فالصحة الجيدة تميل إلى أن تؤدي إلى رضاعة طبيعية ناجحة، وإلى أداء مدرسي متفوق. لذا، فبرغم العلاقة الملاحظة بين هذين العاملين، فإن هذا لا يدعو أن نستنتج بأن مستوى الأداء المدرسي كان النتيجة المباشرة لنوع الرضاعة.

ورغم أن البحث العلّي -- المقارن لا يعد بديلا مقنعا للتجريب، فإنه يوفر مع ذلك طريقة يمكن استخدامها في ظروف يجب فيها إجراء الكثير من البحوث التربوية. وهو يظل طريقة مفيدة يمكن أن تقدم معلومات قيّمة في صنع القرار التربوي.

الشروط اللازمة لاستنباط علاقات علّية

CONDITIONS NECESSARY FOR INFERRING CAUSAL RELATIONSHIP

إذا رغب المرء في الوصول إلى نتيجة مفادها أن متغيراً (X) علّة لمتغير آخر (Y)، فثمة ثلاث بيّنات لازمة:

1- إن علاقة إحصائية بين (X)، (Y) قد أقيمت.
2- إن (X) يسبق (Y) زمنياً.
3- إن عوامل أخرى لم تحدد (Y).

العلاقة الإحصائية هي تلك العلاقة التي يمكن فيها التنبؤ بتغير أحد المتغيرات من خلال تغير في آخر. وعلى أية حال، يجب الملاحظة أن مثل هذه العلاقة بين المتغيرين ليست بحد ذاتها دليلاً كافياً لوجود علّة – و – نتيجة. فيجب على المرء أن يمضي باحثا عن دليل حول المعيارين الآخرين.

وعليه، يجب على الباحث أن يقيم التسلسل الزمني، أي يجب أن يدرس ما إذا كان (Y) قد حدث قبل (X)، ومن ثم لا يمكن أن يكون نتيجة للمتغير (X). فإذا كان (X) علّة (Y)، فإن تغيراً في (X) يجب أن يسبق تغيراً في (Y). أما القرارات حول العلاقة الزمنية بين (X)، (Y) فيمكن إقامتها، إما على أساس منطقي أو نتيجة قياسات تبين أن المجموعات لم تختلف في (Y) قبل تعرضها إلى (X).

إضافة إلى إقامة السلسلة الزمنية، فإن من المهم جداً، أيضاً، أن يدرس الباحث ما إذا كانت ثمة عوامل أخرى غير (X) قد حددت (Y). وهذا هو الدليل الأصعب في الحصول عليه. ويمضي المرء في تدقيق هذا الاحتمال، بإدخال متغيرات أخرى ذات صلة بالتحليل وملاحظة كيف تتأثر العلاقة بين (X)، (Y) بهذه المتغيرات الإضافية. وقد يجد المرء أن العلاقة بين (X) و (Y) تظل صامدة حتى عندما يتم إدخال المتغيرات أخرى. وفي هذه الحالة يجد المرء دليلاً يدعم الاستدلال العلّي. من ناحية أخرى، قد يجد المرء أن متغيرات أخرى قد تؤثر على العلاقة بين (X) و (Y). وفي هذه الحالة يستنتج المرء أن العلاقة بين (X) و (Y) زائفة.

ويقال إن العلاقة زائفة إن لم يكن لمتغيرين أي تأثير على بعضيهما، بل يرتبطان لأن متغيراً آخر يؤثر على كليهما. وبعبارة أخرى، ليس لـ (X) أي تأثير حقيقي على (Y)، بل يوجد بدلا عن

ذلك تأثير متغير دخيل على كل من (X) و (Y). فمثلاً، إن علاقة موجبة بين عدد الكنائس وعدد السرقات المسلحة في المدن في ولاية معينة، لا يعني أن بناء كنائس أكثر سوف يزيد من السرقات المسلحة، ولا زيادة السرقات المسلحة ستسبب بناء كنائس أكثر. والمتغير الدخيل، هنا، هو حجم المدينة. ففي المدن الكبرى كنائس أكثر، وسرقات مسلحة أكثر، بيد أن في المدن الصغيرة سرقات مسلحة أقل وكنائس أقل.

تفسيرات بديلة في البحث العلّي – المقارن

Alternative Explanations In Causal - Comparative Research

الفرق بين المتغير المستقل الفعال والمتغير المستقل المنسوب أمر بالغ الأهمية. فعندما يستطيع الباحثون ضبط المعالجة (X) ثم يلاحظون بعدئذ المتغير التابع (Y)، يكون لديهم دليل معقول بأن (X) يؤثر في (Y). فإن لم يستطيعوا التحكم في / ضبط (X) فقد ينقادون نحو استنتاجات غير مناسبة. فيجب على المرء لدى تقييم البحث العلّي – المقارن، أن يدرس التفسيرات البديلة مثل 1– العلة المشتركة، 2– العلّية المعكوسة، 3– وجود متغيرات مستقلة أخرى.

العلة المشتركة Common Cause

في البحث العلّي – المقارن، يجب على المرء أن يدرس احتمال أن يكون كل من المتغيرين المستقل والتابع لدراسة معينة مجرد نتيجتين منفصلتين لمتغير ثالث. فمثلاً، إذا استخدمنا معدل رواتب المدرسين كمتغير مستقل ومبيعات سلعة كانت ممنوعة كمتغير تابع لكل سنة منذ إلغاء المنع في الولايات المتحدة، لوجدنا ارتبطاً إيجابياً عالياً بين المتغيرين. فهل يعني هذا أنه متى ما ارتفعت رواتب المدرسين فإنهم ينفقون المال على السلعة؟ فالتفسير الأكثر قبولاً هو أن كلاً من رواتب المدرسين ومبيعات تلك السلعة نتيجة للوفرة المتزايدة والتضخم منذ عام 1933.

من الثابت أن معدل دخل خريجي الثانويات الخاصة هو أعلى من معدل دخل خريجي الثانويات العامة. فهل يعني ذلك أن المدارس الخاصة تعد تلاميذها للنجاح المالي بشكل أفضل؟ أم هل أن الفرق الذي يعزى إلى تلك الأسر ذات المال الكافي لإرسال أطفالها إلى مدارس خاصة، هي قادرة على تمويل التدريب المهني لأطفالها أو تنشئ لهم مشروعات في عالم الأعمال أو تشتري لهم مناصب دبلوماسية أو مقاعد في مجلس الشيوخ الأمريكي؟

في مدينة (X)، نجد أن في السنين العشرين الأخيرة، صاحب الزيادة في استهلاك الطاقة الكهربائية زيادة مماثلة في حالات الأمراض العقلية. فهل يعني ذلك أن الزيادة في استخدام الكهرباء يؤدي إلى زيادة في الأمراض العقلية؟ إن تدقيق أعداد الإحصاء السكاني يبين أن سكان مدينة (X) قد ازدادوا عبر السنوات وأن استهلاك الكهرباء وحالات المرض العقلي تنشأ عن النمو السكاني.

ينبغي على الباحث في الدراسات العلّية – المقارنة أن يدرس دائما احتمالات العلّة أو العلل المشتركة التي تتسبب في علاقة مرصودة. وفي أمثلتنا، يمكن بوضوح إلى حد ما تحديد العلل المشتركة. وعلى أية حال، ففي البحث اللاحق للحادث (العلّي – المقارن) يظل دوماً شك مؤرق حول احتمال وجود علل مشتركة لم يفكر أحد بأنها تفسر العلاقة. فلقد اتضح أن معدل إصابات السائقين الذين يستخدمون أحزمة الأمان، أقل من معدل إصابات السائقين الذين لا يستخدمونها. فهل يعزى ذلك إلى أن استخدام أحزمة الأمان يقلل من الإصابات أم هل أن حوادث السائقين الحذرين تؤدي إلى إصابات أقل حيث هم يستخدمون أحزمة الأمان أيضاً؟

العلّية المعكوسة Reverse Causality

في تفسير علاقة ملاحظة / مرصودة في أية دراسة علّية – مقارنة، يجب على المرء دراسة احتمال أن يعزى عكس الفرضية المقترحة إلى الاستنتاج – فبدلا من القول أن (X) يسبب (Y)، ربما الحال هو أن (Y) يسبب (X). فمثلاً، ثمة حقيقة بأن نسبة اتباع الكنيسة الأسقفية البروتستانتية ممن ادرجوا، في موسوعة الأعلام في أمريكا، هي أعظم من نسبتهم في مجموع السكان. فهل يعني ذلك أن الأسقفية البروتستانتية تؤدي إلى نوع من النجاح الذي ينتهي في الأدراج في موسوعة الأعلام؟ إنه أمر ممكن، أو ربما أكثر من ذلك يمكن أن نفترض أن الناس الناجحين يميلون نحو الانجذاب إلى الكنيسة الأسقفية البروتستانتية.

إذا وجدنا أن لدى طلبة الكلية ممن يتناولون طعاماً معيناً، معدل GPA أدنى من الذين لا يتناولون ذلك الطعام، فلا يسعنا أن نستنتج تلقائيا أن استهلاك الطعام يقلل من الأداء الأكاديمي. ربما تدفع العلامات الضعيفة بالطلبة إلى تناول ذلك الطعام. (أو، بالطبع، قد يكون هناك أي عدد من العلل المشتركة التي يمكن أن تقود إلى كل من تناول ذلك الطعام والعلامات الضعيفة).

لقد كشفت البحوث حول تأثيرات ممارسات تربية الأطفال، أن ثمة سلوكاً عدوانياً أكثر لدى الأطفال الذين يخضعون للعقوبة بشكل متكرر. فهل يعني ذلك أن بوسع المرء أن يستنتج أن عقوبة الوالدين تقود إلى أطفال عدوانيين، أم هل أن الأطفال العدوانيين يكونون أكثر عرضة للعقوبة؟

إن فرضية العلّية العكسية اسهل في التعامل من فرضية العلّة المشتركة. فمع الأخيرة قد يكون هناك العديد من العلل المشتركة في كل حالة بما يمكن أن تنشأ عنها علاقة زائفة. ومع العلّية العكسية، هناك فقط احتمال واحد في كل حالة: (Y) سبب في (X) بدلاً من (X) تسبب في (Y).

ففي وضع حيث يسبق (X) دائما (Y) في الزمن، فإن الطبيعة الفعلية لبياناتنا تستبعد احتمال العلّية العكسية. فمثلاً، بينت عدة دراسات أن معدل الدخل السنوي لخريجي الكلية هو أعلى من الدخل السنوي لغير الخريجين. وبوسعنا أن نستبعد فرضية العلّية العكسية، لأن التخرج أو عدمه يسبق الدخل السنوي التالي. إنه لا يسعنا أن نستبعد العديد من العلل المشتركة المحتملة.

وثمة طريقة لإقامة الترتيب الزمني للمتغيرات، تكون بإجراء قياسات لذات الأفراد في

أوقات مختلفة. دعنا نفترض أن أحدهم مهتم بالعلاقة بين قبول المستخدمين لفلسفة شركة والترقية في العمل ضمن تلك الشركة. فإذا أجرى أحدهم مجرد مقابلة لعينة من المستخدمين ووجد أن للذين في المناصب العليا، اتجاهات وأراء تتفق أكثر مع خط النظام القيمي للشركة، فإنه سوف لن يعرف ما إذا كان قبول قيم الشركة وأهدافها سيؤدي إلى الترقية، أم أن الترقية ستزيد القبول بنظام القيم في الشركة. ولاستبعاد العلّية العكسية كتفسير، يمكن للمرء أن يقابل مجموعة متدربين جدد ويحصل عن طريق استبيان، أو مقياس تقديرات، أو ما شابه، على قياس لقبولهم بفلسفة الشركة. وعقب فترة من الزمن، ربما 18 شهراً، يمكن أن يحدد الباحث من سجلات الشركة أياً من المستخدمين قد تمت ترقيته. فإذا دلت الاستنتاجات أن نسبة عالية من الذين كانوا قد عبروا عن اتجاهات وأراء تنسجم مع فلسفة الشركة قد تمت ترقيتهم، مقارنة بمن لم يتم ترقيتهم، فإنه سيكون لدى المرء دليل أفضل بأن التوافق مع فلسفة الشركة قد أدى إلى الترقية. (ويبقى احتمال أن علّة مشتركة أو عدة علل تسبب الفروقات في كل من الفلسفة والترقية).

متغيرات مستقلة محتملة أخرى Other Possible Independent Variables

قد تكون هناك متغيرات مستقلة، عدا المتغير قيد البحث العلّي – المقارن، التي يمكن لها أن تحدث التأثير الملاحظ على المتغير (Y)، أي إضافة إلى (X_1) فإن متغيرات أخرى (X_2) و(X_3) قد تكون عوامل سابقة للتغير في المتغير التابع.

من المعروف أن معدل الانتحار المسجل في السويد هو بين الأعلى في العالم. فهل يعني ذلك أن البيئة السويدية تجعل الناس يلجأون للانتحار؟ هل يعني ذلك أن الشعب السويدي أكثر نزوعاً إلى الانتحار من الشعوب الأخرى؟ وقد تكون هناك حقيقة في إحدى أو كلا هاتين الفرضيتين. ومع ذلك، من الممكن أيضاً أن يكون المتغير المستقل الفعلي هو نزاهة المحققين الجنائيين في السويد مقارنة مع نزاهة المحققين الجنائيين في الأقطار الأخرى. ففي أقطار تلحق وصمة عار اجتماعي، بأسر من ينتحرون، قد يستخدم المحققون الجنائيون كل وسيلة ممكنة لتسجيل الوفاة على أنها بسبب حادث عرضي وليس انتحاراً. وعليه، فالفرق بين معدلات الانتحار المسجلة قد يكون جراء سلوك المحقق الجنائي وليس أي شئ آخر.

ففي مؤتمر لحكام الولايات، يشير الحاكم (X) بفخر إلى معدل الجريمة المنخفض في ولايته. ويشير حاكم آخر، أن قوات الشرطة في ولاية الحاكم (X) قليلة العدد، وأن معدل الجريمة المنخفض قد يشير إلى أن جرائم قليلة جداً فقط، قد تم ذكرها. ويسأل أحد الصناعيين مدير الموظفين لديه عن سبب عدم تعيين خريجي جامعة Tuephingen القديمة، مؤكدا أنه طالما أن العديد منهم يصعدون بسرعة في سلم الترقيات، فإنهم، من الواضح، أكثر كفاءة من الخريجين الآخرين. ويشير مدير هيئة الموظفين بلباقة إلى أن هذه الظاهرة قد لا تتضح عن طريق الكفاءة، بل بحقيقة أن الصناعي نفسه هو خريج تلك الجامعة وهو عن طريق اللاوعي يفضل زملاءه القدامى من خريجي الجامعة في قرارات الترقية.

إن إحدى الواجبات الأولى الواضحة للباحثين هي محاولة تسجيل جميع المتغيرات المستقلة البديلة المحتملة. وبالإبقاء على بقية المتغيرات ثابتة، يمكننا أن نختبر كل متغير لتحديد ما إذا كان يرتبط بـ (Y). فإذا استطعنا أن نستبعد المتغيرات المستقلة البديلة بتفسير كونها غير مرتبطة بـ (Y)، فإنا نحصل على دعم الفرضية الأصلية للعلاقة بين (X) و (Y).

تطبيق لتفسيرات بديلة: في التقرير التالي من الجمعية القومية لمدراء المدارس الثانوية، (1959) توضح البيانات علاقة بين استخدام السيارة والإنجاز / التحصيل الأكاديمي.

هل هناك ارتباط بين السيارات والتحصيل المتميز؟

كلا، تقول ثانوية ماديسون من مدينة روكسبرغ بولاية أيداهو. فقد أوضحت دراسة على معدلات السنوات الأربع لصف ثمائي نموذجي في المدرسة وسائقي السيارات ما يلي:

لم يستخدم تلميذ حاز على A بشكل متواصل، أية سيارة

فقط 15% من الطلبة الذين حصلوا على B، قادوا سيارة إلى المدرسة.

من بين الطلبة (C) 41% جلبوا سيارات إلى المدرسة.

من بين الطلبة (D) 71% قادوا إلى المدرسة.

من بين الطلبة (F) 83% ساقوا إلى المدرسة.

برنامج عمل:

كنتيجة للدراسة، اتخذ مجلس روكسبرغ للتربية قراراً في 11 أب 1958 حدد ما يلي:

1- يجب على الطلبة الجدد وطلبة السنة النهائية في الثانوية ممن يقودون سيارات إلى المدرسة، أن يتقدموا بطلب تحريري لمجلس المدرسة يبين سبب أو حاجة قيادتهم لسياراتهم إلى المدرسة. وسيمنح المرخص لهم تصاريح طلابية.

2- يجب على الطلبة الذين يقودون السيارات بتصاريح طلاب أن يوقفوا السيارات في مناطق مخصصة.

3- سوف تلغى تصاريح الطلبة الذين يخفقون في الإذعان لاتباع القواعد والأنظمة. وإذا أصر الطلبة على عدم الإذعان فسوف يتعرضون للطرد من المدرسة عن طريق مجلس المدرسة.

تعاون المجتمع:

نظم المجتمع برمته دعما لأنظمة مجلس التربية. فقد ادرج المجلس التنفيذي لقانون المواطن ومؤتمر شباب روكسبرغ، مساهمة النادي الأهلي ونادي الروتاري والجمعية الأمريكية والصحف والشرطة والأباء والأمهات في ديناميكية الأمن المروري للمدرسة الثانوية. وقد تعهد مئات

المواطنين بمساندة الدينامية وحصلوا على بطاقة عضوية صادرة عن المجلس والمؤتمر. ويؤيد مدير مدرسة ماديسون، (W.G.Nelson)، بقوة، عمل المجتمع الواسع ضد الاستخدام غير المقيد للسيارات من طلبة الثانوية. "نحن نعتقد أن التقيد باستخدام السيارة سوف يساعد على دوام افضل وانتباه أوثق للدروس. نحن لم "نصل" بعد ولكننا على الطريق".

مدير في تكساس:

في تعليق على برنامج روكسبرغ قال (مدير في تكساس): "كان رأيي منذ وقت طويل أن هناك ارتباط عال بين معدل الرسوب والمنح الدراسية وامتلاك وتشغيل طلبة الثانوية للسيارات بحرية. إننا نجد صعوبة في الحصول على اهتمام كبير من الأولاد ذوي البنية القوية في صفوف الرياضة المتقدمة. لذا، فإن العديد من الأولاد يدفعون للسيارات أو لصيانتها وتشغيلها وذلك بالعمل جزئياً بعد ساعات المدرسة بحيث لا يتسنى لهم الذهاب إلى الرياضة. وإذا توفر لهم مال لتشغيل سياراتهم، فإنهم يفضلون قيادتها حول المدينة بعد المدرسة. وهناك القلة من ذوي الطلبة ممن لهم سيطرة كافية أو تأثير على صغارهم في ردع الاستخدام غير الحكيم لسياراتهم، لكنهم قلة".

هل هناك مسوغ للاستنتاج بأن استخدام السيارة يتسبب في تحصيل أكاديمي منخفض؟

دعنا ندرس الفرضيات البديلة المحتملة؟

1- **علّة مشتركة:** هل هناك متغيرات قد تؤثر على كل من استخدام السيارات والمنح الدراسية؟ نحن نعلم أن الطبقة الاجتماعية مرتبطة بالمنح الدراسية. فإذا كانت للطبقة الاجتماعية تأثير كذلك على السيارة، فإن العلاقة الواضحة بين الدرجات واستخدام السيارة سوف لن تكون علاقة علّة – نتيجة، بل مظهرين لفروقات الطبقات الاجتماعية. إن الفروقات في أسلوب حياة الطلبة أو قيمهم قد تتسبب أيضاً في العلاقة الواضحة. فإذا كان بعض الطلبة يعطون قيمة عالية لقيادة السيارة ولهم اهتمام قليل بالدراسة، فإن عدم إعطائهم سبيلاً للوصول إلى السيارة لن يزيد بالضرورة من دراستهم. وبالوسع اقتراح عدد من الفرضيات العلّية المشتركة القابلة للتصديق.

2- **علّة عكسية:** هل من المحتمل أن تكون الدرجات الضعيفة علّة لاستخدام السيارة؟ يمكننا أن نفترض بشكل معقول أن الطلبة الذين يتدنى أداؤهم في المدرسة، يبحثون عن طرق أخرى للقبول الاجتماعي وأن استخدام السيارة هو أحد هذه الاحتمالات.

3- **متغيرات مستقلة محتملة أخرى:** هل يمكن أن يتصور المدرسون بأن مستخدمي السيارات طلبة لا مبالون فيعطونهم درجات أخفض مما يستحقون؟

هناك العديد من الفرضيات البديلة القابلة للتصديق بما يجعل المرء يتردد في تفسير البيانات على أنها مؤشر لعلاقة علّة – نتيجة. لقد بحث (Asher & Schusler, 1967) القضية ذاتها في

تصميم أقام ضبطاً أكثر من الدراسات السابقة. فبدلا من دراسة الدرجات الحالية فقط واستخدام السيارة، قاما بتسجيل درجات طلبة السنة النهائية في نهاية الفصل الأول ودرجات الطلبة أنفسهم حينما كانوا في سنتهم الأولى. فإذا كان استخدام السيارة يؤثر على الدراسة، فمن المتوقع أن تهبط درجات السائقين بين الوقت الذي كانوا فيه طلبة مستجدين - وعليه غير مؤهلين لإجازات السوق - والوقت الذي كانوا فيه في السنة النهائية. لقد استخدم أشر وشوسلر الفرق بين GPA للطلبة الجدد والنهائيين كمتغير تابع. كما أنهم تباينوا اقترانيا (انظر الفصل 9) في معدل الذكاء. ولم يجدا علاقة دالة (ذات دلالة) بين استخدام السيارة والتغير في الدرجات. لقد استنتجا بأن "حرية الوصول إلى السيارة لا يبدو سببا لهبوط نمط الدرجات أو لصعودها، في هذه الحالة".

الضبط الجزئي في البحث العلّي – المقارن
PARTIAL CONTROL IN CAUSAL - COMPARATIVE RESEARCH

هناك استراتيجيات لتحسين مصداقية البحث العلّي – المقارن، رغم أن أياً منها لا يمكن أن يعوض بصورة كافية عن الضعف الموروث لمثل هذه البحث - أي الافتقار لضبط المتغير المستقل. ومن بين هذه الإستراتيجيات هناك درجات التغير والمزاوجة / التناظر، وتحليل التباين المصاحب/ الاقتراني والارتباط الجزئي والمجموعات المتجانسة ودمج متغيرات دخيلة في التصميم.

درجات التغير Change Scores

في مثال استخدام السيارة، وجدنا أن الاستنتاج الذي تم الوصول إليه، لدى استخدام الفرق بين درجات الطلبة في السنة النهائية ودرجاتهم في السنة الأولى للثانوية، لم يتفق مع الاستنتاج الذي تم التوصل إليه عندما استخدمت الدرجات بذاتها فحسب. وبالتأكيد، فإن للاستنتاج الأول مصداقية أكثر من الثاني لأن الثاني يخفق بصورة كاملة في الأخذ بنظر الاعتبار احتمال أن مستخدمي السيارات كانوا طلبة أفقر من عدم مستخدميها قبل أن يكون لأي منهم سبيل للوصول للسيارات.

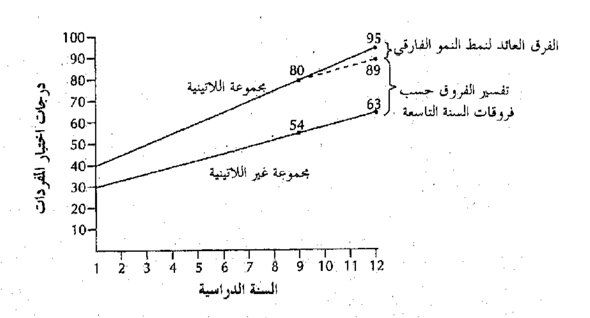
ومع درجات التغير يأخذ المرء بنظر الاعتبار الدرجات السابقة في المتغير (Y) بدلا من الدرجات الحالية في (Y). وحين يقارن المرء مقدار تغير مجموعة المعالجة مقارنة بالمجموعة الضابطة فإنه يكون أقل احتمالا في الوقوع في التضليل مما عليه الحال في استخدام الدرجات الحالية فحسب. ومع ذلك، فإن استخدام درجات التغير حل جزئي فقط ويجب أن تعالج نتائج هذه الدراسات بحذر. مثلاً، ادخل المدير (أ) برنامج قراءة جديد في الصف الخامس، وبعد أن

اصبح قيد الاستخدام لمدة سنة أراد أن يقارن فاعليته مع فاعلية برنامج القراءة الذي حل محله. وفي ذات المنطقة، كانت مدرسة المدير (ب) لا تزال تستخدم البرنامج القديم. وقد أعطت كلا المدرستين اختبار القراءة المقنّن نفسه في نهاية كل سنة دراسية. يقارن المدير (أ) وسط المستوى المكافئ لطلبة الصف الخامس في كلتا المدرستين. وقد وجد أن هذا الوسط هو (6.0) لطلبته بينما وسط طلبة المدير (ب) هو (4.0). فهل هذا دليل مثير لفاعلية الطريقة الجديدة؟ ويدرك المدير (أ) أن الفرق بين الوسطين يمكن أن يعزى إلى الفروقات بين التلاميذ عندما بدأوا الصف الخامس في كلتا المدرستين. وبهذا يحصل على درجات اختبار القراءة الذي اجري حينما كان التلاميذ ينهون الصف الرابع. لقد وجد أن لتلاميذه وسط لمستوى سنتهم الدراسية يكافئ (4.8) بينما حصل تلاميذ المدير (ب) على وسط هو (3.2). وعليه يجب أن يقوم بتعديل إذ لم تكن في الحقيقية كلتا المجموعتين في النقطة نفسها عندما بدأتا الصف الخامس. وبالنسبة لكل تلميذ تيسرت له كلا الدرجتين، طرح المدير (أ) درجة معدل الصف الرابع من درجة الصف الخامس، فوجد وسط فرق يبلغ (1.2+) بالنسبة لتلاميذه و(.8+) لتلاميذ المدير (ب). إن الفرق (.4) بين وسطي درجات التغير لكلا المجموعتين أقل إثارة، لكنه أكثر إقناعا من الفرق (2.0) الناتج عندما استخدمت درجات الصف الخامس فقط ولم يجر أي تعديل على الأداء السابق.

هل بوسع المدير (أ) أن يستنتج الآن أن الطريقة الجديدة هي أكثر فاعلية من القديمة؟ إذا آخذنا بنظر الاعتبار طبيعة درجات التغير فإنه لا يستطيع ذلك. فالفروقات في درجات التغير قد تعزى إلى استمرار الأنماط السابقة. ولأجل إيجاد النمط في هذه الحالة، يقسم المدير (أ) وسط المستوى النهائي المكافئ الحاصل في نهاية السنة الرابعة على أربعة (عدد سنوات الدراسة في المدرسة). وكان لدى تلاميذ المدير (أ) الذين لديهم وسط مكافئ هو (4.8) في نهاية السنة الرابعة، متوسط كسب مكافئ في أداء القراءة هو (1.2) لكل سنة بينما حصل تلاميذ المدير (ب) على متوسط كسب مكافئ في القراءة هو (.8). وفي كلا الحالتين يكون الكسب في درجات القراءة خلال السنة الخامسة هو تماماً ما يتوقعه المرء إذا أخذنا أنماط الكسب السابقة بنظر الاعتبار، وعليه ليس هناك دعم للفرضية القائلة بأن الطريقة الجديدة هي أكثر فاعلية من الطريقة القديمة. ويكمن الضعف الموروث في درجات التغير في أن الكسب الكبير الظاهر في مجموعة، مقارنة بمجموعة أخرى، قد يكون استمرارا لنمط سابق ولا يعزى إلى المعالجة أبدا.

خذ مثالاً آخر. جرى الافتراض بأن دراسة اللاتينية تُحسّن مفردات الإنجليزية لدى طلبة الثانوية. وتوفرت للباحثين درجات المفردات للطلبة في السنتين الأولى والنهائية لكل من اخذوا اللاتينية ومن لم يأخذوها. إذا نظرنا في الشكل (10.1) نرى أن الكسب في درجات المفردات من الصف، 9 حتى الصف 12 بالنسبة إلى مجموعة اللاتينية (من 80 إلى 95 درجة) هي أكبر من كسب المجموعة غير اللاتينية (من 54إلى 63 درجة).

الشكل (10.1): تأثير أنماط نمو المفردات السابقة على درجات التغير



ومع ذلك، فإننا نرى في الشكل 10.1 أن الذين اختاروا اللغة اللاتينية هم مستمرون، فحسب، في نمطهم السابق في زيادة سرعة نمو المفردات وأولئك الذين لم يختاروا اللاتينية يستمرون حسب نمطهم السابق. إن كسب كلا المجموعتين في السنوات الثلاث من الصفوف 9-12 هو نفسه ككسبهم من الصفوف 6-9. وقد يستنتج باحث ساذج ينظر فقط إلى درجات طلبة السنة النهائية لكلا المجموعتين، دون أن يأخذ بنظر الاعتبار أماكن بدئهم المختلفة، بأن الفرق البالغ (32) درجة (63-95) يعود إلى خبرة اللغة اللاتينية. ولعل باحثا أقل سذاجة يستخدم درجات التغير من الصفوف 9 -12 لكلا المجموعتين يعزو فقط (6) درجات (9-15) إلى خبرة اللاتينية. ومع ذلك، فإن هذا الباحث قد توصل أيضاً إلى النتيجة الخطأ. فمن الأكثر احتمالا في إيهام الناس غير الواعين؟ إن استنتاج الباحث الأول هو الأقل احتمالا في القبول، لأن الفكرة القائلة بأن المجموعتين قد تختلفان حول الكفاءة السابقة في مفردات الإنجليزية هي فكرة واضحة إلى حد ما، وأن القرّاء قد يشكون في النتيجة القائمة على درجات طلبة السنة النهائية، فحسب. ويبدو أن استخدام درجات التغير يحل مشكلة الفروقات الموجودة أصلا، واستنتاج الباحث الثاني يبدو أكثر صدقاً. ومع ذلك، فإن استخدام درجات التغير يظل حلاً جزئياً فحسب.

تتكيف درجات التغير في (أ) بدء المجموعات في نقاط مختلفة. وهي لا تتكيف في (ب) الفروقات الموجودة أصلا في معدلات النمو. ومتى ما وجدت الفروقات في معدل النمو فإن الفروقات في درجات ما بعد المعالجة تعزى إلى كل من (أ) و(ب). إن استخدام درجات التغير يزيل التأثير الزائف في (أ) وليس التأثير الزائف في (ب). ومن ثم، فإن درجات التغير تظل حلاً جزئياً فقط. فقد نتوهم بسوء التكيف.

التناظر / المزاوجة Matching

إن الطريقة الشائعة لتامين الضبط الجزئي في البحث العلّي – المقارن هي عن طريق تناظر الأفراد في المجموعتين التجريبية والضابطة حسب أكبر عدد ممكن من المتغيرات الدخيلة. ويتم التناظر عادة على أساس فرد – لفرد وذلك لتشكيل الأزواج المتناظرة. فمثلاً، إذا كان المرء مهتماً بالعلاقة بين خبرات الكشافة والجنوح، فبوسعه تحديد مجموعتين من الأولاد مصنفين على أنهم جانحون وغير جانحين حسب معايير محددة. ومن الحكمة في مثل هذه الدراسة اختيار أزواج من هاتين المجموعتين تتناظر على أساس الوضع الاجتماعي الاقتصادي، والبنية الأسرية، وغيرها من المتغيرات المعروفة بارتباطها بكل من اختيار الخبرة الكشفية والجنوح. ويمكن لتحليل البيانات من العينتين المتناظرتين أن يتم لتحديد ما إذا كانت نسبة الذين شاركوا في الكشافة هي أكبر بين غير الجانحين مما هي بين الجانحين. إن إجراء التناظر في البحث العلّي – المقارن يطرح بعض الصعوبات الموصوفة في نقاشنا حول استخدامه في التجريب في الفصل (9). فمن المقام الأول، يفترض استخدام التناظر في الدراسة العلّية – المقارنة، أن المرء يعرف ما هي العوامل ذات الصلة – أي العوامل التي قد يكون لها ارتباط مع المتغير التابع. وعلاوة على ذلك، يحتمل أن يقلص التناظر عدد الأفراد الذين يمكن استخدامهم في التحليل النهائي. ففقدان الحالات، الموروث في عملية التناظر هو مشكلة حتى أكثر خطورة في البحث العلّي – المقارن مما في التجريب، حيث يسبق التناظر إدخال المتغير المستقل. ففي الأسلوب التجريبي، على الأقل، قد يكون هناك احتمال (ولو أن ذلك لا يوصى به) بإضافة حالات جديدة لتحل محل القديمة التي لا تتناظر. ولا يمكن القيام بذلك في البحث العلّي – المقارن.

وتبين إحدى دراسات (Chapin, 1947) فقدان الأفراد الذي يحدث نتيجة التناظر. لقد كان تشابن مهتماً بتأثير التخرج من الثانوية على نجاح الفرد والتكيف مع المجتمع بعد فترة عشر سنوات. لقد كشفت البيانات الأولية أن الذين اكملوا الدراسة الثانوية كانوا أكثر نجاحا ممن لم يكملوها. ومع ذلك، فإن فحص سجلات الثانوية قد أوضحت أن الذين تخرجوا حصلوا على علامات عالية في المدرسة الابتدائية وكانت أعمارهم أصغر، والمهن الأبوية عالية، وعاشوا في جوار أفضل، وغير ذلك. وربما كانت هذه العوامل عبارة عن عوامل علّية لكل من إكمال المدرسة الثانوية ثم النجاح بعدئذ. ولضبط هذه العلل المشتركة بصورة جزئية، فحص تشابن المجموعات الفرعية للطلبة المتناظرين حسب كل عوامل الخلفية هذه والمختلفين فقط في إكمال المدرسة الثانوية. لقد قلص التناظر عدد الحالات من (1195) إلى (46) أو (23) خريج و(23) غير خريج – أقل من 4% من العدد الذي تمت مقابلته.

ثمة مشكلة اخطر من فقدان الأفراد، وهي دور الارتداد في تصميم الأزواج المتناظرة للعلّي – المقارن. دعنا الآن نستخدم مثال القراءة لتوضح هذه النقطة. فبغية استخدام الأزواج المتناظرة بدلا من درجات التغير غير المتناظرة. فالمدير (أ) سيأخذ طالباً من مدرسته بدرجة قراءة تبلغ

(3.1) في السنة الرابعة ويناظره مع طالب من مدرسة المدير (ب) بدرجة تبلغ (3.1) في السنة الرابعة، وطالبا من مدرسته بدرجة تبلغ (4.8) مع طالب من مدرسة (ب) بدرجة تبلغ (4.8) وهكذا. وعلى أية حال، طالما أن وسط مجتمع (ب) الإحصائي هو أدنى من وسط المجتمع الإحصائي (أ) فسيكون هناك العديد من الطلبة (ب) ذوي الدرجات المتدنية لا يكون لديهم نظير في (أ)، والعديد من الطلبة (أ) ذوي الدرجات العالية ليس لديهم نظير في المجموعة (ب). إن درجات جميع الطلبة غير المتناظرين سوف تستثنى من تحليل البيانات. وبالنسبة لمن يمكن مزاوجتهم / مناظرتهم، فإن الوسط الحسابي لدرجات اختبار القراءة بالنسبة للمجموعتين (أ) و (ب) سيكون متطابقاً. وعليه، يبدو أننا كوّنا بنجاح مجموعة من المدرسة (ب) ممن يتماثلون مع المجموعة من المدرسة (أ) في تحصيل القراءة.

إن هذا كله يبدو جيداً. فهل يمكننا الآن أن نعزو الفروقات في درجات القراءة للصف الخامس إلى الفرق في فاعليتي الطريقتين الجديدة والقديمة؟ كلا، للأسف. إن أزواجنا المتناظرة تجمع أولئك الطلبة ذوي معدلات القراءة الضعيفة في الصف الرابع من المجتمع (أ) وأولئك الطلبة ذوي المعدلات العالية من المجتمع (ب). إن درجات الطلبة (أ) في التناظر سترتد صعودا نحو الوسط الكلي في (أ) وإن درجات الطلبة (ب) في التناظر سوف ترتد نزولا نحو الوسط الكلي في (ب). وهكذا، عندما نقارن درجات القراءة للصف الخامس للمجموعتين المتناظرتين فنتوقع أن يكون الوسط الحسابي (أ) أعلى من الوسط (ب) حتى وإن لم تكن الطريقة الجديدة أكثر فاعلية من القديمة.

تبدو المطابقة جيدة لأنها تقدم مجموعتين تجريبية وضابطة متماثلتين في متغير أو متغيرات ما قبل المعالجة. وعلى أية حال، عندما تستمد مجموعتان متناظرتان من مجتمعين مختلفين، فمن المتوقع أن يخلق الارتداد نحو وسطي المجتمعين الأصليين نتائج زائفة عندما لا يكون المجتمعان متماثلين. فالمزاوجة من داخل مجتمع إحصائي واحد هو في الغالب استراتيجية مفيدة. أما مزاوجة الأفراد من مجتمع إحصائي واحد مع أفراد من مجتمع آخر، فإنها استراتيجية سيئة. وكما هو الحال مع درجات التغير، فإن المزاوجة تتكيف بصورة جزئية مع الفروقات الموجودة أصلا بين المجموعات. وقد يوهمنا التعديل الرديء بالطريقة نفسها التي يوهمنا فيها التعديل / التكيف الرديء في درجات التغير.

تحليل التباين الاقتراني Analysis of Covariance

يمكن استخدام التباين الاقتراني (ANCOVA)، كالتناظر، وذلك لتعديل الفروق جزئياً الموجودة أصلا بين المجموعات. وتكمن فائدة (ANCOVA) في إمكانية استخدام البيانات من جميع الأفراد وليس فقط بيانات الأزواج المتناظرة. لقد رأينا في الفصل (9) أن ANCOVA إجراء ممتاز لتحسين الدقة في تجربة عشوائية. وأيضاً، يُستخدم (ANCOVA) أحيانا، لتعديل

الفروق الأولية، جزئياً، في التصميم العلّي - المقارن. ومع ذلك، طالما كان التعديل في البحث العلّي - المقارن جزئياً، فقط، فإن (ANCOVA) لا يحل مشكلة الفروقات الأولية بين المجموعات، بل يقللها فقط. ولدى تفسير البحث العلّي - المقارن، فإن من غير المناسب الافتراض بأن (ANCOVA) قد عدلت بشكل مقنع الفروقات الأولية.

إن المثال المعروف للمشكلات الموروثة في المطابقة وتحليل التباين الاقتراني هو دراسة (Cicirelli et. al. 1969) العلية - المقارنة لتأثيرات برنامج ميزة البداية(*) (Head Start Program). فقد قارنت هذه الدراسة التحصيل الأكاديمي للطلبة الذين كانوا في برنامج ميزة البداية مع تحصيل من لم يكونوا فيه. وتمت مزاوجة أطفال البرنامج مع من لم يكونوا فيه من المناطق المجاورة وذلك على أساس الجنس (gender)، والمجموعات العرقية، والالتحاق في الروضة. لقد استُخدم تحليل التباين الاقتراني لتعديل الفروقات في الدخل الفردي والمستوى الثقافي للأب والمستوى المهني للأب. ولقد أفادت النتائج أن تحصيل مجموعة "لا ميزة بداية" كان اعظم من مجموعة ميزة البداية حتى بعدما تم تعديل درجات الفروقات الأولية. واستنتج المؤلفون أن تأثير برنامج ميزة البداية كان مؤذياً.

وأشار (Campbell & Erlebacher, 1970) أنه بسبب سوء تعديل، كل من المزاوجة وتحليل التباين الاقتراني، للفروقات الأولية للمجموعات، فإننا نتوقع أن الدرجات المعدلة بعد المعالجة لمجموعة المحرومين، هي أقل من الدرجات المعدلة لمجموعة الأقل حرماناً. وبما أن مدى سوء التعديل غير معروف فإن دراسة سيسيرلي وزملائه لا تمكننا من تحديد ما إذا كانت تجربة ميزة البداية، مؤذية أو مفيدة أو ليس لها أي تأثير.

المجموعات المتجانسة Homogeneous Groups

لعلّك تذكر من النقاش حول الضبط في أجراء التجارب أن من الممكن ضبط تأثيرات متغير معين عن طريق اختيار عينات متجانسة قدر الإمكان حسب ذلك المتغير. ويمكن اتباع أجراء مماثل في البحث العلّي - المقارن. فبدلا من أخذ عينة لا متجانسة ومقارنة المجموعات الجزئية المتناظرة فيها، فإن الباحث قد يضبط متغيراً وذلك بتضمين الأفراد المتجانسين فقط، حسب ذلك المتغير، في العينة. فإذا كان الاستعداد الأكاديمي متغيراً دخيلاً ذا صلة، فبوسع الباحث ضبط تأثيراته عن طريق استخدام أفراد من مستوى استعدادي أكاديمي واحد، فحسب. ويفيد هذا الأجراء في تحرير المتغير المستقل، الذي قد يكون فيه الباحث مهتماً، من متغيرات أخرى يرتبط بها بشكل عام بحيث يمكن لأية تأثيرات توجد أن ترتبط بصورة مسوغة بالمتغير المستقل.

(*) ميزة البداية (head start): مصطلح يشير إلى وجود / إعطاء ميزة لبعضهم في مرحلة مبكرة أو لدى بداية موقف ما. وهو برنامج شرعته الولايات المتحدة لمساعدة الأطفال المحرومين ثقافياً واجتماعياً قبل دخول المدرسة لمضاهاة أقرانهم عند التحاقهم بالمدرسة - (المراجع).

افترض أن أحدا رغب في بحث ما إذا كانت غرفة هادئة في البيت بما فيها من منضدة وكتب وما إلى ذلك، تؤثر الدراسة فيها على درجات طلبة الثانوية. فإذا اختار أحدهم ببساطة مقطعا عرضيا لطلبة الثانوية، وسألهم ما إذا كان لديهم غرفة هادئة يدرسون فيها، ثم قارن درجات من لديهم غرفة مع من ليس لديهم غرفة، فإنه قد يستنتج خطأ أن امتلاك غرفة هادئة للدراسة يؤدي إلى درجات جيدة. أما التفسير البديل فهو إن هناك عوامل أخرى مرتبطة بالطبقة الاجتماعية قد تؤثر على ترتيبات الدراسية والدرجات التي يحصل عليها الطلبة في المدرسة. فمثلاً، قد تكون الأسر ذات الدخل المتدني مكتظة بالأفراد، واحتمال ما تؤكده على التحصيل الأكاديمي أقل مما تفعله الأسر في الطبقة الوسطى أو العليا. فإن أراد أحدهم ضبط تأثيرات الطبقة الاجتماعية فيمكنه أن يحدد الدراسة بأفراد من مستوى اجتماعي واقتصادي واحد، فحسب. وإن وجد أحدهم علاقة بين ترتيبات الدراسة والدرجات ضمن طبقة اجتماعية واحدة فسيكون لديه ثقة أكثر إلى حد ما، في الاستنتاج بأن الفرق في الدرجات يعزى إلى ترتيبات الدراسة وليس إلى الفروقات الاجتماعية والاقتصادية. وبالطبع، فإن هذا الإجراء يحدد تعميم الاستنتاجات على طبقة اجتماعية واحدة استخدمت في الدراسة، فلا يعلم المرء ما إذا كانت العلاقة قائمة في الطبقات الاجتماعية الأخرى.

يعد استخدام العينات المتجانسة حلا جزئيا فقط للمشكلات المتأصلة في البحوث العلّية – المقارنة. فبوسعنا ضبط بعض متغيرات العلّية المشتركة عن طريق اختيار عينات متشابهة في متغير علّي مشترك مشكوك فيه. ولا يمكننا التأكد أن لدينا أفراداً متجانسين في جميع المتغيرات العلّية المشتركة المشكوكة وغير المشكوكة. ففي مثالنا، يمكن أن يكون امتلاك أو عدم امتلاك غرفة هادئة وتحصيل أكاديمي داخل أسر الطبقة الوسطى، معتمداً على درجة تقييم العائلة للثقافة، أو الحجم العائلي، أو متغيرات أخرى. ويمكن أن يكون الارتداد مشكلة أيضاً مع العينات المتجانسة. فمثلاً، إذا استخدم المدير(أ) في مثالنا السابق فقط أطفالاً في المدرستين أحرزوا درجات بلغت (4.0) بالضبط في اختيار القراءة للصف الرابع، فإن الارتداد سيجعل مجموعة (أ) ترتد صعودا ومجموعة (ب) ترتد نزولا تماما كما حدث فعلا عندما تناظر الأفراد.

الارتباط الجزئي، وهو طريقة أخرى لتعديل الفروقات الأولية في الفصل 11.

دمج متغيرات دخيلة في التصميم

Building Extraneous Variables Into The Design

قد يكون ممكناً دمج متغيرات مستقلة دخيلة ذات صلة، في التصميم العلّي – المقارن، واستخدام أسلوب تحليل التباين العاملي. مثلاً، افترض أن الذكاء هو متغير دخيل ذو صلة وأنه ليس من الممكن ضبطه من خلال أية وسيلة أخرى. في مثل هذه الحالة، يمكن إضافة الذكاء إلى التصميم كمتغير مستقل آخر، ويصنف أفراد الدراسة في إطار مستويات الذكاء. ويتم تحليل قياسات

المتغير التابع من خلال تحليل التباين، وقد يتم عندئذ تحديد التأثيرات الرئيسية والتفاعلية للذكاء. سيكشف مثل هذا الإجراء أية فروقات دالة بين المجموعات في المتغير التابع، دون افتراض أية علاقة علّية بين الذكاء والمتغير التابع. وقد تكون هناك متغيرات أخرى دخيلة تعمل لتوليد كل من التأثير الرئيسي وأي تأثير تفاعلي.

مثلاً، تذكر العديد من الدراسات علاقة بين الترتيب الولادي والتحصيل الثقافي والإنجاز المهني: فمن يولد أولاً يعمل افضل من الذي يولد ثانياً، والذي يولد ثانياً يعمل أفضل من الثالث وهكذا. وبعد مراجعة عميقة للأدبيات استنتج (Schooler, 1972) بأن: "معظم ما تكرر ذكره حول الفروقات بين مراتب الولادة – التكرار الأكبر لأوائل المواليد بين مجموعات اتسمت بالتحصيل الفكري غير الاعتيادي، أو الإنجاز المهني – تبدو الأكثر شحاً / ندرة في تفسير دلالة الفروقات بين توجهات الطبقة الاجتماعية في حجم الأسرة". ولأن هناك علاقة عكسية بين الطبقة الاجتماعية وحجم الأسرة فإن متوسط مكانة الطبقة الاجتماعية للأطفال الخوامس(*) سيكون أدنى من متوسط مكانة الطبقة الاجتماعية للأطفال الروابع، وهكذا. وبما أن الطبقة الاجتماعية مرتبطة بالتحصيل، فإن العلاقة الواضحة بين الترتيب الولادي والتحصيل، قد يعزى فقط إلى الفروقات في متوسط مكانة الطبقة الاجتماعية وليس إلى ترتيب نسق الولادة بذاته.

ولاحقاً، درس (Belmont & Morolla, 1973) الترتيب الولادي ودرجات اختبار (Raven) للمصفوفات المتتابعة (وهو اختبار ذكاء غير شفوي) لقرابة (400,000) شاب في نذرلاندز. وقد وجدا أنه ضمن كل حجم أسرة، هبط الذكاء مع الترتيب الولادي. ونظراً لأنهما كانا يقارنان الدرجات فقط ضمن أسرة الخمسة أطفال وضمن أسرة الأربعة أطفال، وهكذا، فإن الفروقات في الطبقة الاجتماعية لا يمكن أن تعزى إلى نتائجهما.

هل تمكننا نتائج بلومنت ومورولا الآن من القول أن الترتيب الولادي بذاته مرتبط بالذكاء؟ وسيكون من الحكمة معالجة هذا فقط كنتيجة مؤقتة. وأحد أسباب تحفظنا، هو أنه قد تكون هناك متغيرات، عدا الطبقة الاجتماعية، تعزى إلى تلك العلاقة الواضحة. فبالتأكيد، إن متوسط العمر لأبوي الأطفال الروابع هو أكبر من متوسط عمر أبوي الأطفال الثوالث ولعله يكون لهذا شيء ما ذو صلة بالظاهرة. وربما يعزى ذلك إلى شئ لا يسعنا حتى أن نتصوره.

إن دمج / بناء متغيرات أخرى في التصميم العلّي – المقارن هو حل جزئي، لكنه لا يمكننا أن نتأكد بأننا اخترنا المتغيرات الصحيحة، أو إننا استخدمنا "جميع" المتغيرات التي كان ينبغي دراستها.

لقد تم التأكيد في هذا الفصل، على أن الضعف الرئيسي للتصميمات العلّية – والمقارنة

(*) الأطفال الخوامس (Fifth Children) هم الأطفال الذين يكون كل منهم خامس مولود في الأسرة – (المراجع)

يكمن في الافتقار إلى الضبط. ونظر لأنه ليس من الممكن استخدام العشوائية لتخصيص الأفراد في المجموعات أو القيام بتفعيل مباشر للمتغير المستقل في وضع تم ضبطه، فإن هناك إمكانية دائما لوجود متغيرات لم تضبط تكون مسؤولة عن التباينات في المتغير التابع. وبسبب ذلك، لا يمكن للمرء الافتراض أن المجموعات كانت متماثلة في بداية الدراسة. ولأنه ليس للباحث إمكانية ضبط من تعرض للخبرة ومن لم يتعرض لها، فإنه من الممكن أن يكون هناك شيء آخر، يعود للناس أو لبيئتهم، يحدد التعرض في المقام الأول. وعليه، فقد يكون هناك "الشيء الآخر" غير الخبرة بذاتها، هو الذي يكون المتغير المستقل الحاسم.

كنتيجة للضبط غير الكافي في الدراسة العلّية – المقارنة، فإن تفسير الاستنتاجات قد يكون متوقفاً على مخاطرة بشكل خاص. أن خطر التفسير غير الصحيح كبير. فعندما تقام علاقة بين متغيرين في إطار دراسة علّية – مقارنة، فإن التحليل يكون، إلى حد ما، قد اتخذ أولى خطواته. ويجب على الباحث أن يدرس ويختبر أية فرضيات بديلة ممكنة، وحتى بعد القيام بذلك، يجب عليه أن يدرك بأن أية علاقة واضحة قد تعزى إلى علة يتعذر فهمها. فإجراءات معينة، مثل استخدام درجات التغير، والمزاوجة، وتحليل التباين الاقتراني، والارتباط الجزئي، والمجموعات المتجانسة، ودمج متغيرات دخيلة في دراسة معينة يمكن لها أن تساعد على تجنب أخطاء جسيمة في الدراسات العلّية – المقارنة، إلا أنها جميعا تعجز عن تعديل الفروقات القبلية للمعالجة بين المجموعات. إن إدراج فرضيات بديلة (علّة مشتركة، علّية عكسية، ومتغيرات مستقلة متبادلة) يمكن لها أن تساعد على تقييم النتائج العلّية – المقارنة بشكل أكثر واقعية.

تصميم البحث العلّي – المقارن

DESIGNING THE CAUSAL - COMPARATIVE INVESTIGATION

يوضح الشكل التالي التصميمين الأساسيين للبحث العلّي – المقارن:

	متغيرات مستقلة		متغيرات تابعة
التصميم 1	أفراد معروفون باختلافهم في المتغير المستقل	←	يختبر الباحث فرضية تخص المتغير التابع المحتمل
التصميم 2	يختبر الباحث فرضية تخص المتغير المستقل المحتمل	→	أفراد معروفون باختلافهم في المتغير التابع

Design 1 التصميم1

إن أحد الأمثلة للتصميم 1 هو دراسة مقارنة أداء حل مشكلة لطلبة كلية مبدعين وغير مبدعين.

وتقول الفرضية: سيبدى طلبة الكلية المبدعون سرعة أكبر ودقة في واجب حل مشكلة أكثر من الطلبة غير المبدعين. وتوضح هذه الفرضية، الحاجة للتصميم العلّي – المقارن، لأنه ليس في وسع الباحثين تفعيل الإبداع ولا تخصيص الطلبة للمجموعات بصورة عشوائية. فيجب أن يبدأوا بمجموعتين تختلفان مسبقاً في المتغير المستقل، وهو الإبداع، وتقارنان حسب المتغير التابع وهو أداء حل المشكلة.

يجب أن يعرف الباحثون "طالب الكلية المبدع" و "طالب الكلية غير المبدع" حسب تعريفات إجرائية دقيقة. فقد يجرى تعريف طلبة الكلية المبدعين بكونهم الطلبة الذين يحصلون على درجات فوق الربع المئيني الثالث في كل من اختبار Guilford للاستخدامات المتبادلة ونتائجها، واختبار الجناس التصحيفي (تغير أحرف كلمة لتشكيل كلمة جديدة – المترجم). فالطلبة الذين درجاتهم دون الربع المئيني الأول في الاختبارات سيعرفون بأنهم غير مبدعين.

ينبغي أن يحاول الباحثون تحديد متغيرات، عدا الإبداع، التي يمكن لها أن تؤثر على المتغير التابع وهو أداء حل المشكلة واتخاذ الخطوات لمساواة المجموعة التجريبية والضابطة في هذه المتغيرات بالمزاوجة أو بالوسائل الإحصائية. فمثلاً، في هذه الدراسة ثمة متغيرات أخرى ينبغي ضبطها هي الذكاء، والجنس، وربما موضوع الاختصاص في الكلية أو سنة الكلية. فقد يتناظر طالب ذكي في المجموعة المبدعة مع مثيله في المجموعة غير المبدعة.

بعد تشكيل المجموعتين المتناظرتين، تعطى كلا المجموعتين مقياسا للمتغير تابع وهو واجب حل مشكلة. وسيكشف تحليل للبيانات عن طريق اختبار– t أية فروقات دالة في أداء حل المشكلات لكلا المجموعتين وربما يبين علاقة بين الإبداع وأداء حل المشكلات. ورغم أن المرء قد يستنتج من دراسة كهذه أن هناك علاقة، إلا انه لا يستطيع أن يفترض صلة علّية بين الإبداع وأداء حل المشكلات. وقد تكون هناك متغيرات أخرى لم تضبط يمكن لها أن تؤثر بشكل منفرد أو بتوليفة منها على حل المشكلات. ويدرك الباحث اليقظ، الحاجة لفحص التفسيرات الأخرى البديلة للاستنتاج العلّي-المقارن.

ويكون المتغير المستقل أحيانا، منسوباً، لا يمكن تفعيله – كالجنس أو الحال الاجتماعي الاقتصادي أو العرقية. وتعتبر الدراسات المشمولة بتحليل ما بعد التحليل في (Hyde, Fennema, & Lamon, 1990) لأداء الرياضيات لدى الذكور والإناث (المذكور في الفصل 5) أمثلة عن دراسات من هذا النوع.

ويكون المتغير المستقل أحيانا متغيراً يمكن تفعيله إلا أنه، نموذجياً، يتعذر ضبط الباحث له لأسباب عملية أو أخلاقية. مثلاً، ليس أخلاقياً، تخصيص بعض الطلبة عشوائيا لحجزهم في الروضة، بينما تتم ترقية طلبة آخرين كي تتم دراسة تأثيرات الحجز على الأداء التالي. وعلى أية حال، فإن الفهم العميق لهذه القضية يمكن تحقيقه عن طريق مزاوجة الطلبة الذين تم حجزهم مع

الذين لم يتم حجزهم ، حسب أكبر عدد ممكن من المتغيرات مثلاً، زاوج (Mantzicopoulos & Marrison, 1992) الطلبة الذين تم الاحتفاظ بهم في الروضة مع الطلبة في ذات المدرسة الذين لم يحتفظ بهم حسب متغيرات الجنس، والعمر، والوضع الحرج، وتحصيل القراءة، وتحصيل الرياضيات. لقد قارنا الأطفال الذين تم الاحتفاظ بهم مع الأطفال من نفس العمر الذين تم ترقيتهم إلى مستوى أعلى من المجموعة المحتفظ بها، ثم مع أطفال أصغر بسنة من السنة الدراسة نفسها. ودلت النتائج على فائدة لصالح الأطفال المحتفظ بهم خلال سنتهم الثانية في الروضة، ولكن ليس أبعد من ذلك. واستنتج المؤلفان أن نتائجهما لا توحي بأن الاستبقاء / الحجز سياسة مؤثرة على الصغار ذوي الوضع الحرج.

التصميم2 2 Design

في التصميم 2، وهو النوع العلّي المقارن الثاني، تختلف مجموعتان أو أكثر في متغير تابع ويتم اختبار فرضيات تخص متغيرات مستقلة محتملة. خذ عشرة أعضاء من فريق مصارعة يصيب ثلاثة منهم مرض عند العودة من دورة للمباريات. ولبحث علّة المرض تسأل الطبيبة ماذا أكلوا حين توقف الفريق لتناول الغداء وهم في طريق عودهم إلى البيت. وتجد الطبيبة أن كل بند اختاره المصارعون المرضى الثلاثة كان مختلفا باستثناء الدجاج بالكريمة. أما المصارعون السبعة الذين لم يتناولوا الدجاج بالكريمة فلم يمرضوا. وهذا يقودها إلى الاستنتاج بأن الدجاج بالكريمة هو علّة المرض. وبهذا تستخدم الطبيبة طريقة الاتفاق والاختلاف المشتركة للفيلسوف جون ستيوارت ميل (John Stuart Mill):

> "إذا حدثت الظاهرة في حالتين أو أكثر وكان لها ظرف واحد مشترك، بيد أنها لم تحدث في حالتين أو أكثر لا تشترك بشيء، عدا غياب ذلك الظرف، فإن الظرف وحده الذي تختلف فيه مجموعتا الحالات، هو المؤثر، أو العلّة، أو الجزء الضروري من العلّة للظاهرة" .(1846, p. 229)

وكمثال آخر لتصميم البحث العلّي – المقارن الثاني، هو دراسة (Benn, 1968) عن متغيرات ذات صلة بضمان الرابطة بين الأبناء بعمر 18 شهرا وأمهاتهم العاملات. وقد تم تقييم نوعية رابطة (متغير تابع) عن طريق تحليل أشرطة فيديو للأولاد في أحداث مصورة ومصممة لتقييم توازن الرابطة والسلوك الاستكشافي في وجود وغياب الأم وشخص غير مألوف. وقد وجد أن بين الأولاد الذين طمأنوا للرابطة هذه، اتسمت العلاقة بين الأم – الابن بالدفء، والقبول، وحرية التعبير العاطفي. وعادت أمهات الأولاد، المطمئنين لرباطهم بأمهاتهم، إلى العمل مبكرات خلال السنة الأولى من عمر الطفل الصغير أكثر من أمهات الأولاد الذين لم يطمئنوا لتلك العلاقة. ولم تكن للطبقة الاجتماعية الاقتصادية، ونوع الرعاية المقدمة للطفل، صلة بالرابطة بين الأم وطفلها.

دور البحث العلّي – المقارن

THE ROLE OF CAUSAL - COMPARATIVE RESEARCH

إذا أخذنا بنظر الاعتبار المخاطر التي ينطوي عليها البحث العلّي – المقارن لوجدنا أن هناك العديد من أبناء المهنة ممن يقولون بأنه لا يجب أن نشغل أنفسنا بهذا النوع من البحث على الإطلاق. واقتناعهم أساساً، يقوم على أن الاعتراف بالجهل، أفضل من المجازفة بالتوصل إلى استنتاجات غير صحيحة.

من ناحية أخرى، هناك الذين يرون أن العديد من المتغيرات المهمة جداً لنا، ليست سهلة التطبيق في البحث التجريبي. فليس بوسعنا أن نخصص بصورة عشوائية أطفالا لأسر مفككة أو سليمة، وإلى طبقة اجتماعية عالية أو متدنية، وإلى جماعات الأقران ذوي التوجه التحصيلي أو عدمه، وما إلى ذلك. وعليه، إذا أردنا أن نتعلم شيئاً عن علاقات بين مثل هذه المتغيرات المنسوبة وغيرها من المتغيرات فإن أسلوب العلّية – المقارنة هو سبيلنا الوحيد. فإذا استخدمنا الطرق المناسبة لضبط جزئي وأخذنا بنظر الاعتبار فرضيات بديلة فربما يكون بوسعنا أن تكون على صواب أكثر مما نكون على خطأ.

هناك، بالتأكيد، العديد من الدراسات العلّية المقارنة تتمتع بثقة عالية. فدراسة كبير الأطباء عن العلاقة بين التدخين وسرطان الرئة هو مثال معروف. فليس من الممكن تخصيص مجموعة من الأفراد عشوائيا كي يدخنوا لعدة سنوات ومجموعة أخرى لا تدخن، ومن ثم يجب إجراء الدراسة كبحث علّي – مقارن. إن فرضية العلّية العكسية بأن سرطان الرئة يجعل الناس يدخنون ليست ممكنة. ولم يبد أن أياً من فرضيات العلّة المشتركة قابل للتصديق: الناس عصبيو المزاج يكونون عرضة لكل من التدخين وسرطان الرئة، وأن بعض الإستعداد الجيني يقود إلى كليهما، وهكذا. لقد ضبط كبير الأطباء العديد من المتغيرات المستقلة البديلة – مثلاً، بتحليل العينات بصورة منفصلة عن مناطق التلوث الجوي العالي والتلوث الجوي المتدني. فالأدلة التجريبية مع الحيوانات التي حُملت على استنشاق دخان السيجارة وأخرى لم تستنشقه تمخض بدوره عن دليل على علاقة علّة بنتيجة. وإذا أخذنا كل ذلك بنظر الاعتبار، فإنه رغم المخاطر المتأصلة في البحث العلّي – المقارن فسوف يستنتج الجميع أن المراهنة على أن هناك علاقة علّة بنتيجة بين التدخين وسرطان الرئة بين البشر، أفضل من المراهنة بعدم وجود مثل هذه العلاقة.

لقد لوحظ تأثير كثير من البحوث العلّية المقارنة على التربية. فمتغيرات مثل الخلفية البيتية، والتبعة الجينية، وتلف الدماغ، والخبرات المبكرة هي متغيرات تربوية مهمة جداً حتى وإن كانت تتجاوز ضبط التربويين لها.

إن بوسعنا جميعا شجب العديد من حالات ، شواهد استخدام البحث العلّي – المقارن في أوضاع كان يمكن فيها استخدام التصميمات التجريبية أو على الأقل، شبه التجريبية. ولعل ما

يُفزع هنا هو عدد المرات التي قامت بها الوكالات المحلية أو الرسمية والحكومة الفدرالية "باختبار" برامج جديدة عن طريق إدخال جميع الأفراد المؤهلين إلى البرنامج ثم محاولة تقييم الأثر اللاحق للبرنامج. وغالباً ما استجابت الوكالات الحكومية، بما في ذلك المدارس إلى الضغط بالقيام "بشيء ما" بخصوص مشكلة معينة وذلك بتصميم برنامج جديد لجميع الأفراد المؤهلين. وقد وصف Evans (من مكتب التربية الأمريكي) و Schiiler (من مكتب الفرصة الاقتصادية) (1970) الطريقة التي تستجيب بها الوكالات الحكومية لهذا الضغط:

"سوف تستمر محاولات تنفيذ الشرط المطلوب للتخصيص العشوائي، في مواجهة اعتراضات زبائن البرنامج على أساس أن مثل هذه الإجراءات تنطوي على حرمان اعتباطي من البرنامج بالنسبة إلى الذين خصصوا للمجموعة الضابطة. ومن بين المستائين، سيجهر الصُرحاء منهم بالشكوى لدى المسؤولين ورجال الكونغرس. ونتيجة لذلك، فإن مدراء البرنامج سوف يرغبون في تجنب هذا الإجراء ليكونوا إلى جانب المعارضين له.

وتقودنا خبرتنا إلى الاستنتاج، ولو بالرغم منا، أنه في الظروف الزمنية الفعلية الضاغطة المعبأة سياسياً، التي تبرز فيها برامج العمل الاجتماعي بصورة حتمية، فإن الحالات التي يتم فيها التخصيص العشوائي نادرة، وأن طبيعة العمليات السياسية والحكومية تجعل من المحتمل استمرار ذلك الوضع. ولسوء الحظ، فإن العملية السياسية ليست منظمة، أو مبرمجة أو منطقية. وتنطلق موجة من الجمهور والكونغرس تأييداً لبرامج العمل الاجتماعي، حيث تتضخم غالباً بسرعة مع قليل من التوقع. وحينما يتم سنّ التشريع، فإن الضغوط رجال الإدارة، للقيام بتنفيذ برنامج سريع، تكون شديدة. وفي مثل هذه الظروف – التي تعد القاعدة وليس الاستثناء – فإن الالتماسات الداعية إلى ضرورة تنفيذ البرنامج بحرص، في مساق خطة عامة لتجربة حقيقية مع تخصيص عشوائي للأفراد بحيث يتسنى لنا بثقة أن نقيّم فاعلية البرنامج، يكون مصيرها التجاهل".

إنا على اقتناع بأن أي استياء يواجهه الزبائن "المحرومون" من البرنامج الجديد ليس إلا قطرة، مقارنة بفيض استياء من دافعي الضرائب ممن يكتشفون أن الملايين قد أنفقت على برامج افتقرت إلى طريقة مخططة بشكل جيد لتحديد ما إذا كانت هذه البرامج قد أنجزت شيئاً بالفعل. فلابد من وجود طرق لمعالجة مشكلات العلاقات العامة للتخصيص العشوائي. أليس بوسع الحكومة أن تقدم للعديد من المدارس برنامجاً مموّلاً بشكل تام لنصف تلاميذها بشكل عشوائي؟ وإذا كان البرنامج جذابا، فإنا نعتقد أن العديد من الجماعات ستكون راغبة في المساهمة بمثل هذه التجربة، باعتبار أن نصف رغيف عشوائي أفضل من لا رغيف على الإطلاق . وعندما يتوفر الدليل، فبوسع الحكومة أن تقدم لجميع التلاميذ تلك البرامج التي اتضحت قيمتها وتُسقط ما لا قيمة له.

الخلاصة SUMMARY

يستخدم البحث العلّي - المقارن عندما لا يكون الباحثون في وضع يختبرون فيه فرضية معينة عن طريق تخصيص أفراد حسب ظروف مختلفة يفعّلون فيها بشكل مباشر المتغير المستقل. ففي البحث العلّي - المقارن تكون التغيرات في المتغير المستقل قد حدثت مسبقاً، وتحتم على الباحثين دراستها باستعادة الأحداث من أجل آثارها المحتملة على متغير تابع ملاحظ / مرصود.

ورغم أن هناك العديد من المساوئ في أسلوب العلّية - المقارنة، إلا إنها مع ذلك، نظل غالباً الطريقة الوحيدة التي يستطيع بواسطتها الباحثون التربويون الحصول على المعلومات الضرورية حول ميزات مجموعات محددة من الطلبة، أو المعلومات اللازمة للصياغة البارعة للبرامج في المدرسة. إنها تتيح للباحثين استقصاء الأوضاع التي يستحيل فيها إدخال تباين موجه / قابل للضبط. وثمة صفات مميزة مثل الاستعداد الأكاديمي، والإبداع، والوضع الاجتماعي الاقتصادي، وشخصية المدرس، لا يمكن تفعيلها، ومن ثم يجب بحثها من خلال البحث العلّي - المقارن وليس من خلال الطريقة التجريبية الأدق.

إن احتمال العلاقات الزائفة قائم دائماً في البحث العلّي - المقارن. فبأخذ احتمالات العلّة المشتركة، العلّية المعكوسة والمتغيرات المستقلة البديلة، بعين الاعتبار، قد يساعدنا على تقييم مثل هذه البحوث بشكل أكثر واقعية. وهناك عدة إستراتيجيات للضبط الجزئي يمكن لها أن تساعدنا على تجنب الأخطاء الكبيرة في التصميمات العلّية - المقارنة، إلا أن أحدا لن يكون بوسعه أن يحل المشكلات المتأصلة، كلياً، في هذه التصميمات. وعلينا أن نمارس الحذر دائما في تفسير نتائج العلّية - المقارنة.

مفاهيم أساسية Key Concepts

active independent variable	متغير مستقل فعال ، نشيط
analysis of covariance (ANCOVA)	تحليل التباين الاقتراني
attribute independent variable	متغير مستقل منسوب
building extraneous variables into a design	إدماج / دمج متغيرات دخيلة في تصميم
causal relationship	علاقة عليّة
change score	درجة تغير
causal-comparative research	بحث علّي - مقارن
common cause	علّة مشتركة
ex post facto research	بحث لاحق / راجع للحادث
homogeneous groups	مجموعات متجانسة

matching	مزاوجة / تناظر
other independent variables	متغيرات مستقلة أخرى
partial control	ضبط جزئي
reverse causality	علّية معكوسة / عكسية
spurious result	نتيجة زائفة / وهمية

تمارين EXERCISES

1- كيف تختلف المتغيرات المنسوبة والمتغيرات الفعالة / النشيطة عن بعضها؟

2- تحت أية ظروف يمكن للمرء أن يستخدم البحث العلّي – المقارن؟ وما هي نقطة الضعف الرئيسة للتصميمات العلّية – المقارنة؟

3- ما هي الشروط اللازمة لاستنباط أن (X) أحدثت (Y)؟

4- أي من الفرضيات البحثية التالية تدعو إلى البحث التجريبي وأي منها تدعو إلى البحث العلّي – المقارن؟

أ- لا يحتفظ الصغار بقصة قرأها غريب لهم مدة زمنية تعادل مثيلتها في حالة قراءة أمهاهم للقصة ذاهما.

ب- دافع التحصيل لدى طلبة الصف الخامس المبدعين أعلى مما لدى الطلبة غير المبدعين.

ج- للمراهقين من أسر أحد الوالدين سجلات في الشرطة أكثر من المراهقين لأسر كلا الوالدين.

د- سيحصل طلبة الصف الأول ممن تعلموا التهجئة صوتيا على درجات أعلى في اختبار التهجئة ممن تعلموا التهجئة باستخدام طريقة الكلمة الكاملة.

هـ- طلبة الجامعة الذين يحصلون على درجة أعلى من 1200 في اختبار سجل التخرج (GRE) سيحصلون على درجات أعلى في كلية التربية من الطلبة الذين درجاهم لم تتجاوز 1200.

و- مفهوم الذات لدى الأطفال المعاقين بدنيا أدنى من الأطفال غير المعاقين.

5- عرف، وأعط مثالا لكل من المصطلحات التالية

أ- العلّة المشتركة.

ب- العلّية المعكوسة / العكسية.
ج- المتغيرات المستقلة الأخرى .

6- ما الذي تتكيف معه درجات التغير، وما الذي لا تتكيف معه؟

7- ما هي فوائد ومساوئ المزاوجة / التناظر؟

8- كيف يمكن للباحث أن يتعامل مع متغير دخيل ذي صلة لا يمكن ضبطه من خلال المزاوجة أو الوسائل الأخرى؟

الأجوبة ANSWERS

1- المتغير المنسوب هو سمة يتملكها الفرد قبل بدء الدراسة، لذا لا يمكن تفعيله بشكل مباشر. أما المتغير الفعال / النشيط فهو متغير مستقل، غير أنه يمكن للباحث أن يفعّله (يحدث تغييراً عليه) بشكل مباشر.

2- يستخدم البحث العلّي – المقارن حين تكون هناك متغيرات مستقلة منسوبة – أي حيث يمتلك المرء المتغيرات المستقلة قبل بدء الدراسة، ولذلك فإن هذه المتغيرات غير قابلة للتفعيل. أما الضعف الرئيس للتصميمات العلّية – المقارنة فهو الضبط. وبما أنه تستحيل العشوائية وتفعيل المتغيرات المستقلة فإن المتغيرات غير المضبوطة قد تكون مسؤولة عن التباين في المتغير التابع.

3- يجب إقامة علاقة إحصائية بين (X) و (Y): يجب أن تسبق (X)، (Y) في الزمن، ويجب أن يكون هناك دليل بأن العوامل الأخرى لا تحدد (Y).

4- أ) التجريبية، ب) العلّية – المقارنة، ج) العلّية – المقارنة،
د) التجريبية، هـ) العلّية – المقارنة، و) العلّية – المقارنة.

5- أ- تعني "العلّة المشتركة"، أنه إذا كان متغيران (س) و(ص) على صلة ببعضيهما، فمن الممكن ألا يكون أي مهما علّة للأخر، إلا أن كليهما قد حدثا بفعل متغير ثالث (ع). فمثلاً، إذا اتضح أن الدرجات العالية في اختيار تحصيل ودرجات عالية في عمل أكاديمي على صلة ببعضيهما، فإن من الممكن أن تكون علّة كليهما عامل ثالث وهو الاستعداد العام.

ب- تعني "العلّية العكسية" أن عكس الفرضية المقترحة يمكن أن يتسبب في العلاقة. فمثلاً، قد يلاحظ المرء أن هناك علاقة بين البطالة والإفراط في النوم، ويفترض أن الإفراط في

النوم يجعل العامل يفقد عمله، بينما في الواقع قد يكون من يفقدون عملهم يميلون إلى الإفراط في النوم.

ج- "المتغيرات المستقلة الأخرى"، هي متغيرات غير التي جرت ملاحظتها، هي التي قد تكون مسؤولة عن العلاقات. فمثلاً، من المعروف أن نسبة الناس المحجوزين في المستشفيات العقلية بسبب انفصام الشخصية، هي أكبر في الولايات المتحدة مما في بريطانيا، بينما نسبة الناس المحجوزين في بريطانيا بسبب الكآبة هي أكبر من النسبة في أمريكا. وهذا لا يعني بالضرورة أن البيئة الأمريكية هي مؤدية أكثر إلى انفصام الشخصية وأن البيئة البريطانية تؤدي أكثر إلى الكآبة. فقد يكون مرد ذلك إلى أن البريطانيين أكثر تسامحا مع أعراض انفصام الشخصية، لذا فهم أقل احتمالاً في حجز هؤلاء الذين يعانون من هذه الأعراض، بينما الأمريكيون هم أكثر تسامحاً مع ظهور أعراض الكآبة.

6- تتكيف درجات التغير مع الفروق الأولية بين المجموعات. وهي لا تتكيف مع أنماط التغير الفارقية.

7- تقدم المزاوجة مجموعات متكافئة حسب متغير خاص أو مجموعة متغيرات قبل المعالجة. ويمكن توقع أن يشوه الارتداد الإحصائي النتائج حين تكون المجموعات المتناظرة من مجتمعات إحصائية مختلفة. وتقلص المزاوجة عدد الأفراد المتيسرين.

8- قد يدمج الباحث متغيراً دخيلاً ذا صلة، في تصميم علّي - مقارن، ثم يستخدم أسلوب تحليل التباين. ويتطلب هذا الأجراء تحليل الآثار الرئيسية والتفاعلية.

REFERENCES المصادر

Asher, W., and M.M. Schusler (1967). Students' grades and access to cars. *Journal of Educational Research,* 60, 10.

Belmont, L., and Marolla, F.A. (1973). Birth order, family size, and intelligence. *Science,* 182, 1096-1101.

Benn, R.K. (1986). Factors promoting secure attachment relationships between employed mothers and their sons. *Child Development,* 57, 1224-1231.

Campbell, D. T., and Erlebacher, A. (1970). How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In V. Hellmuth (Ed.), *Compensatory education: A national debate: Vol.* 3. *Disadvantaged child.* New York: Brunner/Mazel.

Chapin, F.S. (1947). *Experimental designs in sociological research,* New York:

Harper & Row.

Cicirelli, V., et al. (1969). *The impact of Head Start: An evaluation of the effects of Head Start on childrens cognitive and affective development.* (A report presented to the Office of Economic Opportunity pursuant to contract B89-4536.) Westinghouse Learning Corporation and Ohio University. (Distributed by the Clearinghouse for Federal Scientific and Technical Information, U.S. Department of Commerce, National Bureau of Standards, Institute for Applied Technology. PB 184328.)

Evans, V.W., and Schiller, J. (1970). How preoccupation with possible regression artifacts can lead to a faulty strategy for the evaluation of social action pro- grams: A reply to Campbell and Erlebacher. In V. Hellmuth (Ed.), *Compensatory education: A national debate: Vol. 3. Disadvantaged child.* New York: Brunner/ Mazel.

Mantzicopoulos, P., and Morrison, D. (1992). Kindergarten retention: Academic and behavioral outcomes through the end of second grade. *American Educational Research Journal,* 29, 182-198.

Mill, J.S. (1846). *A system of logic.* New York: Harper & Brothers.

National Association of Secondary School Principals (1959). Do automobiles and scholarship mix? *NASSP Spotlight on Junior and Senior High Schools,* 36, 3.

Rogerson, B.C.F., and Rogerson, C.H. (1939). Feeding in infancy and subsequent psychological difficulties. *Journal of Mental Science,* 85, 1163-1182.

Schooler, C. (1972). Birth order effects: Not here, not now! *Psychological Bulletin,72,* 161-175.

الفصل الحادي عشر

البحث الارتباطي
Correlational Research

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون بوسع الطالب أن:

1- يصف طبيعة البحث الارتباطي.

2- يصف الطرق التي يستخدم فيها البحث الارتباطي.

3- يصف تصميم البحث الارتباطي.

4- يدرج الأنواع المختلفة لمعاملات الارتباط ويذكر شروط استخدامها المناسب.

5- يفسر معاملات الارتباط في إطار الحجم والدلالة الإحصائية والأهمية العملية.

6- يختبر فرضية حول معامل الارتباط (R).

7- يعرف عامل التنبؤ والمعيار/ العامل المتنبأ به (المتغير التابع).

8- يطور معادلة الارتداد للتنبؤ من بيانات معينة حول متغير واحد (Y) من متغير آخر (X).

9- يعرف الارتداد المتعدد ويوضح متى يستخدم.

10- يوضح ظاهرة الانكماش في الارتداد المتعدد.

11- يعرف التحليل المميز/ التمييزي ويوضح متى يستخدم.

12- يعرف التحليل العاملي ويوضح هدفه.

13- يوضح دور الارتباط في التحليل العاملي.

14- يعرف الارتباط الجزئي ويوضح هدفه.

15- يصف دراسة بحثية يكون فيها الارتباط الممتد مناسباً.

16- يعرف تحليل المسار ويوضح المسار وهدفه باختصار.

يهتم البحث الارتباطي بتحديد العلاقات القائمة بين المتغيرات. وليس هناك تفعيل للمتغيرات – لا أكثر من استقصاء مدى ارتباط المتغيرات. وحين تكون هناك علاقة بين متغيرين نقول إنهما مرتبطان. وتوصف قوة العلاقة واتجاهها عن طريق مؤشر كمي يدعى "معامل الارتباط" – تذكر من الفصل (5) أنه لدى تفسير معامل الارتباط، ينظر المرء إلى كل من إشارته وحجمه. فإشارة معامل الارتباط (+ أو –) توضح اتجاه العلاقة. فإذا كان هذا المعامل موجبا، فهذا يعني انه كلما ازداد متغير يزداد المتغير الأخر. مثلاً، الارتباط بين الارتفاع والوزن موجب لأن الناس طوال القامة يميلون إلى أن يكونوا أثقل وزناً، بينما قصار القامة أخف وزناً. ويوضح المعامل السالب انه عندما يزداد متغير يتناقص المتغير الآخر. فمعامل الارتباط بين درجة الحرارة الخارجية خلال اشهر الشتاء وفواتير التدفئة سالبة.

ويشير حجم معامل الارتباط إلى قوة العلاقة بين المتغيرات. فالمعامل يمكن أن يتراوح في القيمة من (1.00+) (موضحا العلاقة الموجبة تماما) إلى (0) (موضحا عدم وجود علاقة) حتى يصل إلى (1.00-) (موضحا العلاقة السالبة تماما). وتعني العلاقة الموجبة تماما أن لكل زيادة مقدارها وحدة معيارية (Z) في متغير واحد، يزداد المتغير الآخر بوحدة معيارية (Z) متناسبة معها. وتشير العلاقة السالبة تماماً إلى أن كل زيادة أحد المتغيرين بوحدة يقابله نقص في الآخر بوحدة متناسبة معها، ويبين القليل جداً من المتغيرات الارتباط التام، خاصة في ارتباط السمات الإنسانية. ويتراوح أغلبها ما بين (0) و ($\pm$1.00) مثل (40.+) للارتباط بين مقياس الاستعداد الدراسي ومعدل النقاط في الثانوية، أو (40.-) للارتباط بين أيام الغياب عن المدرسة ودرجة اختبار التهجئة. أما الارتباط بين درجات الاستعداد والوزن لعينه من البالغين فيتوقع أن تكون صفراً.

استخدامات الارتباط USES OF CORRELATION

البحث الارتباطي مفيد في العديد من الدراسات. ويمكن تصنيف التطبيقات الرئيسية للارتباط حسب: 1) التنبؤ، 2) تقييم الثبات، 3) وصف العلاقات.

التنبؤ Prediction

رأينا في الفصل (8) أن الارتباط يستخدم لتقديم دليل متصل / مرتبط بالمعيار لصدق الاختبار. ويوضح معامل الارتباط مدى ارتباط / صلة درجات الاختبار بمقاييس / بوحدات المعيار المعاصرة مثل مدى ارتباط درجات اختبار تهجئة مع تقديرات التهجئة الراهنة في بطاقات تقارير التلاميذ (الصدق التلازمي) أو مع درجات المعيار المستقبلي مثل تقديرات التهجئة في بطاقات التقارير بعد سنة (صدق تنبؤي).

لا يتحدد تقييم الفائدة التنبؤية بالاختبارات. فبوسعنا أيضا بحث مسائل مثل مدى ارتباط الوضع الاجتماعي الاقتصادي بالتحصيل المدرسي الراهن والنجاح المهني المستقبلي، وفائدة التقديرات المدرسية في الثانوية وتوصيات المدرسين في التنبؤ بمعدل الفصل الدراسي في الكلية (GPA) أو العلاقة بين سنوات خبرة المدرسين وتقديرات الجدارة التي يحددها المدير لهم.

إن تقييم الفائدة التنبؤية للعديد من الإجراءات مهم على نحو خاص عندما تستخدم هذه التنبؤات لصنع قرارات مهمة حول الأشخاص، كتحديد مستواهم في شعب للصف قد تكون علاجية، أو عادية، أو متقدمة، أو القبول أو عدم القبول في الكلية.

تقييم الثبات Assessing Consistency

رأينا في الفصل (8) أن ثبات (اتساق) أي اختبار يمكن تقييمه من خلال ارتباط الاختبار – إعادة الاختبار أو الأشكال / الصور المتكافئة أو درجات التجزئة النصفية. ويمكن استخدام الارتباط لقياس الثبات (أو افتقاره) في العديد من الحالات المختلفة. مثلاً، ما مدى ثبات تقديرات الجدارة التي يحددها المدير ونائبه بشكل مستقل للمدرسين في مدرسة؟ وما مقدار الاتفاق بين الحكام الأولمبيين في تقدير أداء مجموعة من لاعبي الجمباز؟ وحين يطلب من مجموعة مدرسين ترتيب حدة الفوضى التي يولدها كل بند من قائمة بالاضطرابات السلوكية، فما مدى اتفاق ترتيباتهم؟

وصف العلاقات Describing Relationships

تساعد طرق البحث الارتباطي على توضيح العلاقات والأنماط بين المتغيرات. مثلاً، قد تستخدم الدراسات العليّة – المقارنة إجراءات ارتباطيه للإجابة على أسئلة مثل: ما هي العلاقة بين الوضع الصفي(جديد أم متقدم) وتسامح طلبة الكلية مع أساليب الحياة البديلة؟ وما هي العلاقة بين قلق الطلبة والأداء في اختبار موقوت(*) ؟

(*) اختبار موقوت (Timed test): هو الاختبار المألوف حيث ينبغي على الطالب الإجابة على مجموعة الأسئلة في زمن محدد – (المراجع).

وتكون الدراسة الارتباطية غالباً، استكشافية أساساً – فالباحث يسعى إلى تحديد نمط العلاقات الموجودة بين متغيرين أو أكثر. فالمعلومات التي توفرها مثل هذه الدراسات الارتباطية مفيدة بشكل خاص حين يحاول المرء إدراك مفهوم بنائي معقد، أو بناء نظرية حول ظاهرة سلوكية ما. ويجب أن يتذكر الباحث، على أية حال، بأنه ينبغي أن تستند المتغيرات المنتخبة لدراسة من هذا النوع، إلى نظرية، أو إلى بحث سابق، أو إلى ملاحظات الباحث. فطريقة "بندقية الصيد" حيث يجرى المرء ارتباطات لعدد من المتغيرات لمجرد تفحص ما قد يبرز على السطح، هي طريقة لا يوصي بها. وقد يكون الباحث في بعض الدراسات الارتباطية قادرا على صياغة فرضية حول العلاقة المتوقعة. مثلاً، من النظرية الظاهراتية(*) قد يفترض المرء أن هناك علاقة موجبة بين إدراك أطفال السنة الأولى لذواتهم وتحصيلهم في القراءة. وفي أمثلة أخرى قد يفتقر الباحث إلى المعلومات الضرورية لصياغة فرضية.

وتوضح دراسة (Bourke, 1986) استخدام الطريقة الارتباطية في بحث العلاقة بين حجم الصف وتحصيل الطلبة. لقد افترض أن تنفيذ ممارسات التدريس المختلفة في الصفوف ذات الأحجام المختلفة تؤدي إلى اختلافات في التحصيل. وكانت الخطوة الأولى لبورك هي استقصاء الارتباطات البسيطة بين حجم الصف وبعض الممارسات التدريسية. ويبن الجدول 11.1المتغيرات التي لها ارتباطات مميزة مع حجم الصف.

الجدول 11.1: علاقات متغيرات ممارسة التدريس مع حجم الصف

المتغير	الارتباط مع حجم الصف
استخدام تدريس مجموع الصف	-.24
عدد المجموعات المستخدمة في الصف	.33
مدرس يتفاعل مباشرة مع الطلبة	-.27
مقدار الضوضاء المسموح بها	.34
مدرس يدقق في الأفكار بعد السؤال	-.25
أسئلة الطلبة	.32
مدرس ينتظر الاستجابة	-.20
الواجب البيتي وتقييم التعيينات	-.28
الإدارة غير الأكاديمية	.29

(*) الظاهراتية (phenomenology): فرع من الفلسفة يهتم بالشعور / الوعي والفكرة والخبرة – (المراجع).

يبين الجدول 11.1 أن هناك أسئلة أكثر لطلبة في الصفوف الأكبر حجما (عادة يبحثون عن مساعدة أو توضيح) غير أن متابعة المدرس للأسئلة كانت أكثر في الصفوف الأصغر حجماً. وكان، نموذجياً، الاستخدام الأكبر للواجبات البيتية في الصفوف الأصغر. أما مقدار الزمن الذي ينفقه المدرسون في التفاعل المباشر مع الطلبة فهو أكبر في الصفوف الصغيرة. ورغم أن الدراسة هي أكثر عمقا مما يوضحه المثال، فأنه بالوسع رؤية كيف يبدأ الباحث بارتباطات لمساعدته في فهم ظاهرة معقدة، وهي في هذه الحالة التفسيرات الممكنة للعلاقة بين حجم الصف والتحصيل. وقد تستخدم إجراءات ارتباطيه معقدة أخرى (كالتحليل العاملي) لدراسة العلاقات التبادلية المتزامنة للعديد من المتغيرات وذلك لتوفير فهم أعمق لمكونات مفاهيم بنائية مثل الاتجاهات نحو المدرسة، وأسلوب القيادة، والتفكير التجريدي.

تصميم الدراسات الارتباطية
DESIGN OF CORRELATIONAL STUDEIES

التصميم الأساسي للبحث الارتباطي سهل ويسير التنفيذ. أولاً، يحدد الباحث المتغيرات المهمة ويطرح سؤالاً عن العلاقة بينهما. ويتم تحديد المجتمع الإحصائي في ذلك الوقت. مثلاً، قد يكون سؤال البحث الارتباطي هو: ما هي العلاقة بين القدرة الكمية (قدرة التعامل مع الكم/ الأعداد- المراجع) وتحصيل العلوم بين طلبة المدارس الثانوية؟

وبعدئذ يحدد الباحث كيف يتم إضفاء سمة الكمية / التكميم على المفاهيم البنائية كالقدرة والتحصيل في المثال أعلاه. وقد يدرك سلفاً التعريفات الإجرائية المقبولة للمفاهيم البنائية كما قد يبحث عن تعريفات في مصادر كتلك الموصوفة في الفصل (3)، أو قد يطور تعريفاته ثم يقيم ثباتها وصدقها. وفي المثال، قد يقرر الباحث أن القدرة الكمية سوف تتحدد كدرجات في اختبار القدرة في المدرسة والكلية، السلسلة III (SCAT III) كما يجرى تعريف التحصيل العلمي كدرجات في الشعب العلمية للاختبارات المتتابعة للتقدم التربوي)STEP III).

من المهم في الدراسات الارتباطية اختيار أو تطوير مقاييس تكون مؤشرات مناسبة للمفاهيم البنائية قيد البحث، ومن المهم بشكل خاص أن يكون لهذه الأدوات ثبات مقنع وأن تكون صادقة لقياس المفاهيم البنائية قيد الدرس . ويتأثر حجم معامل الارتباط بكفاية أدوات القياس لتحقيق هدفها المقصود. مثلاً سوف لن تميز الأدوات التي تكون سهلة جداً أو صعبة جداً لأفراد الدراسة، فيما بينهم وسوف تؤدي إلى معامل ارتباط اصغر من الأدوات التي لها مستوي صعوبة مناسب. ولا يحتمل أن يكون للأدوات التي ثباتها متدن وصدقها مشكوك فيه، نتائج مفيدة.

بعد ذلك، ينتخب الباحث العينة لتمثيل المجتمع الإحصائي المحدد. ونموذجياً، لا تتطلب

الدراسات الارتباطية عينات كبيرة جدا . فبالوسع افتراض أنه إذا كانت علاقة قائمة، فإنها ستكون واضحة في عينة ذات حجم معتدل (مثلاً 100-50). ولا نوصي بعينات أقل من (30) فردا. وإذا كان الباحث مهتماً بتعميم الاستنتاجات على مجتمع إحصائي كبير، فإنه سيحتاج إلى سحب عينة عشوائية من المجتمع الإحصائي . فمثلاً، إذا أراد باحث تعميم النتائج من دراسة ارتباطيه للقدرة في الرياضيات والتحصيل العلمي على جميع طلبة الثانوية في منطقة، فإنه يجب أن يسحب عينة عشوائية من بين جميع طلبة المدارس الثانوية في المقاطعة.

وأخيراً، يجمع الباحث البيانات الكمية حسب متغيرين أو اكثر للعينة ثم يحسب معاملات الارتباط بين الدرجات.

معاملات الارتباط Correlation Coefficients

ثمة عدة أنواع مختلفة من معاملات الارتباط . ويختار المرء الإحصاءة المناسبة على أساس سُلم/ ميزان القياسات للمقاييس المستخدمة وعدد المتغيرات .

معامل ارتباط بيرسون التتابعي

Pearson's Product Moment Coefficient of Correlation

لقد تعرفت في الفصل (5) إلى معامل ارتباط بيرسون التتابعي، الإحصاءة الوصفية الأوسع استخداما للارتباط. وقد تذكر أن معامل ارتباط بيرسون مناسب للاستخدام عندما تتوزع المتغيرات، المراد دراسة ارتباطها، بصورة اعتيادية/ نظامية وتقاس حسب سلم قياس / مقياس فتري أو نسبي.

معامل الارتباط "رو" لسبيرمان

Spearman's Rho Coefficient of Correlation

في البحث، نرغب أحيانا إيجاد معامل الارتباط بين مجموعتين من القياسات التي تتبع نظام الرتب (الترتيب التدريجي) أي بيانات ترتيب وليس فترة. مثلاً، قد نريد ربط الرتب التي عينها أثنان من المدرسين لمجموعة من الطلبة فيما يخص الأصالة. والمؤشر المستخدم في مثل هذه الحالات هو معامل ارتباط سبيرمان "رو" (rho) ورمزه "ρ" الذي يحسب عن طريق المعادلة التالية:

$$\rho = 1 - \frac{6\Sigma D^2}{N(N^2-1)} \qquad (11.1)$$

حيث

ρ = معامل ارتباط سبيرمان rho

ΣD^2 = مجموع مربعات الفروقات بين الرتب

N=عدد الحالات .

وللإيضاح أنظر الجدول (11.2) الذي يبين ترتيب 11 تلميذا من اثنين من المدرسين. يعرض العمودان (2)، (3) من الجدول (11.2) ترتيب المدرس (1) والمدرس (2) على التعاقب. أما العمود (4) فيبين الفروقات بين هذه الرتب. مثلاً، الفرق بين ترتيب التلميذ (أ) من قبل هذين المدرسين هو (3-) والتلميذ (ب) هو (1-) وهكذا. أما مجموع القيم في هذا العمود فهو دائما (0). ويعطي العمود (5) مربعات هذه الفروقات . أن مجموع قيم (D^2) هو (26) وعدد الحالات هو (11). وعندما تعوض هذه القيم في المعادلة (11.1) فإن الحساب يعطي معامل سبيرمان لارتباط الرتب وهو (+.88).

وعند ترتيب الأفراد أو الأشياء بهدف إيجاد معامل ارتباط بين مجموعتين من الرتب، فمن المحتمل أن تجد لأثنين أو اكثر الرتبة ذاتها. مثلاً، عندما يكون لأثنين من الأفراد الرتبة (3) فإنهما سيكونان الثالث والرابع في السلسلة ومن الضروري تخصيص الموقع المتوسط - في هذه الحالة (3.5). أما الشخص التالي فسيخصص للمرتبة (5).

ويحتاج المرء أحيانا أن يجد العلاقة بين جملة رتب ومجموعة قياسات فترية، مثل مجموعة درجات اختبار. وليس ممكناً الارتقاء بالرتب من بيانات ترتيب إلى بيانات فترات، فعلى المرء أولاً، تحويل الدرجات إلى رتب ثم يطبق معادلة "رو" لسبيرمان.

جدول 11.2: حساب معامل الارتباط بين مجموعتين من الرتب

(1) الطالب	(2) رتبة المدرس الأول R_1	(3) رتبة المدرس الثاني R_2	(4) الفرق D	(5) D^2
أ	1	4	-3	9
ب	2	3	-1	1
ج	3	1	+2	4
د	4	2	+2	4
هـ	5	5	0	0
و	6	6	0	0
ز	7	8	-1	1
ح	8	9	-1	1
ط	9	7	+2	4
ي	10	11	-1	1
ك	11	10	+1	1
			0	26

$$\rho = 1 - \frac{(6)(26)}{11(121-1)} = +.88$$

أنظر، مثلاً في الجدول 11.3. فالعمود (2) من الجدول يوضح ترتيب المدرس لمجموعة طلبة حسب استعدادهم في الرياضيات. ويبين العمود (3) الدرجات الفعلية لهؤلاء الطلبة في اختبار الاستعداد بالرياضيات.

ولتحديد العلاقة بين الرتب والدرجات الفعلية، فعلى المدرس أن يحوّل درجات الطالب في الاختبار إلى رتب. العمود (4) يبين رتب الطلبة حسب درجاتهم في الاختبار. لاحظ أن كلا من لمياء ودرويش قد أحرز (17) درجة، لذا ارتبطا بالمرتبتين الثالثة والرابعة، وعليه فقد أعطيا مرتبة متساوية هي (3.5) التي تعد الوسط للرتبتين (3) و(4). والوضع مماثل بالنسبة إلى توفيق وجمال وداود الذين أحرز كل منهم (15) درجة وتقاسموا المراتب الخامسة والسادسة والسابعة. المرتبة (6) التي تعد المعدل للمراتب (5) (6) (7) تخصص لكل من الثلاثة. أن إجراء إيجاد قيم (D) و (D^2) يشبه بالضبط الإجراء المبين في الجدول (11.2). أما معامل الارتباط بين هذه المجموعات من البيانات فهو (+.95).

الجدول 11.3 حساب ارتباط الرتب بين مجموعة بيانات ترتيب ومجموعة بيانات فترة

(1) الطالب	(2) رتبة 1	(3) الدرجات	(4) رتبة 2	(5) D	(5) D^2
جميل	1	19	1	0	0
لمياء	2	17	3.5	-1.5	2.25
لمى	3	18	2	+1	1
درويش	4	17	3.5	+0.5	0.25
توفيق	5	15	6	-1	1
منال	6	14	8	-2	4
جمال	7	15	6	+1	1
داود	8	15	6	+2	4
جيهان	9	12	10	-1	1
أمل	10	13	9	+1	1
جاسم	11	8	11	0	0
سحر	12	5	12	0	0
				0	15.5

$$\rho = 1 - \frac{(6)(15.5)}{12(144-1)} = +.95$$

يعد ارتباط سبيرمان للرتب حالة خاصة من الارتباط التتابعي لبيرسون وتشتق معادلته من معادلة بيرسون (r). وهكذا، فمن الممكن استخدام المعادلة (5.15) لحساب الارتباط بين المجموعتين من الرتب. فمثلاً لو طبقنا المعادلة (5.15) على رتب الجدول (11.2) فسوف نحصل على معامل ارتباط مماثل هو (+.88). واستثناءا لذلك، هو حالة الرتب المتعادلة / المتساوية. فمع الرتب المتعادلة كتلك الموجودة في الجدول (11.3) سوف لا تؤدي معادلة بيرسون إلى معامل ارتباط مماثل للمعامل الذي حصلنا عليه بتطبيق معادلة سبيرمان للرتب.

ويعد معامل ارتباط سبيرمان للرتب جزءاً من العائلة الإحصائية كالوسيط. فهو إحصاءة ترتيب مصمم للاستخدام مع بيانات ترتيب. ومثل معامل ارتباط بيرسون التتابعي، فإنه يتراوح من (-1.00) إلى (+1.00). وحيث يكون لكل فرد الرتبة ذاتها في كلا المتغيرين، فإن الارتباط (rho) سيكون (+1.00)، وعندما تكون الرتب في متغير هي المعاكسة تماما لها في المتغير الآخر فإن (rho) سيكون (-1.00). وإذا لم تكن ثمة علاقة بين الرتب فإن معامل ارتباط المرتبة سيكون (0). ويفسر (rho) سبيرمان بطريقة بيرسون (r) نفسها.

مؤشرات أخرى للارتباط Other Indexes of Correlation

إضافة إلى مؤشرات بيرسون وسبيرمان هناك عدة مؤشرات أخرى للارتباط مناسبة لإيجاد قوة العلاقة بين الأنواع المختلفة للمتغيرات. وتعرض الصفحات التالية بعض هذه المؤشرات الأخرى وتطبيقاتها – ولكن، دون الدخول في تفاصيل حسابية. وينصح الطلبة المهتمون بمراجعة الكتب الإحصائية لهذه الإجراءات الحسابية.

الارتباط الثنائي والارتباط الثنائي الأصيل Biserial and Point Biseral Correlation. جاء تطوير أسلوبي الارتباط الثنائي والارتباط الأصيل لتلبية حاجة المرء. مثلاً، في إيجاد العلاقة بين متغير فترة أو نسبي ومتغير اسمي ثنائي التفرع. ويتطلب استخدام الارتباط الثنائي افتراضاً ينص على أن المتغير الذي يقع تحت التفرغ الثنائي هو متغير مستمر/ متصل واعتيادي. وبعبارة أخرى يتكون التفرع الثنائي بصورة مصطنعة من متغير مستمر. مثلاً، قد يعمل المرء على تصنيف الأفراد بكونهم دون المتوسط وأعلى من المتوسط في مجال الإبداع على أساس درجاتهم في اختبار للإبداع. ولو ربطت درجات اختبار الذكاء (المتغير المستمر) مع الإبداع الذي قيس على أنه فوق ودون المتوسط (التفرع الثنائي المصطنع) فإن معامل الارتباط الثنائي سيكون مقياسا مناسبا للعلاقة.

أما الارتباط الثنائي الأصيل، فيستخدم عندما يكون التفرع الثنائي أصيلا – مثلاً عند دراسة العلاقة بين الجنس ودرجات في اختبار استدلالي. وتخصص للتفرع الثنائي قيم عددية من (1) أو (0) وتترابط هذه القيم مع قيم المتغير المستمر. وكمتغيرات ثنائية التفرع هي المواطن / اللا مواطن، الأعسر / الأيمن، الخريج / غير المتخرج، المدخن / غير المدخن وما إلى ذلك. إن الارتباط الثنائي الأصيل هو تبسيط رياضي لمعامل بيرســـون (r).

الارتباط الرباعي ومعامل فاي (Phi) The Tetrachoric Correlation and the Phi Correlation. يستخدم الارتباط الرباعي ومعامل (فاي) لإيجاد العلاقة بين المتغيرات عندما يكون كلاهما ثنائي التفرع. وتحديد طبيعة التفرع الثنائي المؤشر الخاص المراد استخدامه. وإذا كان كل من المتغيرين مستمرا فعلاً، إلا الهما متفرعان بصورة مصطنعة، يكون معامل الارتباط الرباعي مؤشراً مناسباً. وان أراد أحدهم إيجاد الارتباط بين تفرعين ثنائيين فعلاً، فإن معامل ارتباط فاي هو المؤشر المستخدم. ويستخدم الارتباط الرباعي مثلاً لإيجاد العلاقة بين الإبداع والذكاء حينما يكون قد تم تفريع كل من المتغيرين ثنائياً إلى قياسات دون المتوسط وفوق المتوسط. وهو ليس واسع الاستخدام مؤشراً ارتباط. ونفقد معلومات كثيرة بتقليص المتغيرات المستمرة إلى تفرعات ثنائية. ويطبق معامل فاي "Phi" بغية وصف العلاقة بين الجنس لطلبة الثانوية الاقدمين وما إذا كانوا قد فازوا بمنحة للدراسة الجامعية. ويتفرع الجنس ثنائيا إلى ذكر – أنثى وأن "الفوز بمنحة الدراسة الجامعية" هو "نعم – كلا" وكلاهما تفرعان ثنائيان اصيلان. وتخصص القيم العددية (1، 0) إلى التفرغات الثنائية ويتم حساب معامل "Phi". ويعتبر معامل فاي كذلك تبسيطا حسابيا لمعامل بيرسون (r). ويلخص الجدول 11.4 معاملات الارتباط المستخدمة مع أنواع مختلفة من المقاييس/ السلالم.

الجدول 11.4 أنواع معاملات الارتباط وأنواع المقاييس المطابقة

معامل الارتباط	نوع المقياس/ السلّم
التتابعي – بيرسون	سلم/ مقياس فترة أو نسبي، سمة لكلا المتغيرين.
الرتب – سبيرمان	مقياس ترتيبي، سمة لكلا المتغيرين.
ثنائي أصيل	متغير واحد في مقياس فتري، المتغير الاخر هو متغير أصيل ثنائي التفرع على مقياس اسمي.
ثنائي	متغير واحد على مقياس فترة أو نسبي، الآخر تفرع ثنائي مصطنع(*).
رباعي	ثنائي التفرع مصطنع (مقياس اسمي) يستخدم كلا المتغيرين. ولكليهما توزيعات مستمرة.
معامل "فاي– Phi"	ثنائي التفرع حقيقي (مقياس اسمي) سمة لكلا المتغيرين.

نسبة الارتباط: إن الافتراض الذي ينطوي عليه معامل الارتباط التتابعي هو أن العلاقة بين المتغيرين خطية. وتكون هذه العلاقة أحيانا خطية منحنية. أي توصف العلاقة بخط منحني وليس

(*) التفرع الثقنائي المصطنع (artificial dichotomy): يشير إلى تقسيم متغير مستمر / متصل، كيفيا، إلى صنفين– مثلا، تقسيم درجات اختبار إلى صنفين – نجاح، رسوب – باستخدام نقطة فاصلة. أما مثال التفرع الثنائي الحقيقي فهو ذكر – أنثى، حي – ميت.

بمستقيم. وفي مثل هذه الحالات لا تكون معادلة بيرسون (r) مناسبة لأنها تؤدي إلى بخس تقدير درجة العلاقة. ويستخدم مؤشر آخر وهو نسبة الارتباط (ايتا eta) عندما تكون العلاقة بين المتغيرين (X) و (Y) منحنية. وبالنسبة للعلاقات غير الخطية فإن من المناسب اكثر تقييم قوة الارتباط بين المتغيرات.

افترض مثلاً أن أحدهم أراد إيجاد الارتباط بين العمر والقوة البدنية. إنا نعلم أن القوة البدنية مرتبطة إيجابيا بالعمر، إلى حد معين، تكون بعده الزيادة في العمر مصحوبة بنقصان في القوة البدنية. إن نسبة الارتباط سيكون مؤشرا مناسبا يمكن استخدامه في هذه الحالة.

الارتباط الجزئي والمتعدد Partial and Multiple Correlation

إن أساليب الارتباط التي نوقشت حتى الآن مناسبة للاستخدام مع متغيرين فقط. وهناك أوضاع يتحتم على المرء فيها أن يتعامل مع اكثر من متغيرين. و"الارتباط الجزئي" هو أسلوب يستخدم لتحديد أي ارتباط يبقى بين المتغيرين عندما يتم إقصاء تأثير متغير آخر. والارتباط بين متغيرين قد يحدث لان كليهما مرتبط بمتغير ثالث. والارتباط الجزئي يضبط هذا المتغير الثالث. مثلاً، افترض أن أحدهم كان مهتماً بالارتباط بين المفردات ومهارات حل المشكلة. فكلا هذان المتغيران يرتبطان بمتغير ثالث وهو العمر الزمني. مثلاً، الأطفال بعمر الثانية عشرة، لديهم مفردات متطورة الأكثر من الأطفال بعمر الثامنة، ولديهم مهارات حل المسائل اكثر تطورا أيضا. فدرجات المفردات ومهارات حل المشكلات سوف ترتبط مع بعضها، لأن كليهما يرتبطان بالعمر الزمني. ويستخدم الارتباط الجزئي مع بيانات كهذه لغرض الحصول على قياس ارتباط مع استبعاد تأثير العمر. أما الارتباط الباقي بين المتغيرين عندما يستبعد ارتباطهما مع متغير ثالث فيدعى بالارتباط الجزئي من المرتبة الأولى. ويمكن استخدام الارتباط الجزئي لإزالة تأثير اكثر من متغير. وعلى أية حال، نظراً لصعوبة التفسير فإن الارتباط الجزئي الذي ينطوي على إقصاء اكثر من متغير لا يستخدم في الغالب.

وينطوي "الارتباط المتعدد" أيضاً على أكثر من متغيرين. ويمكّن هذا الأسلوب، الباحثين من إيجاد أفضل ترجيح ممكن لمتغيرين مستقلين أو اكثر للحصول على أقصى ارتباط مع متغير تابع واحد. مثلاً، استخدم الباحثون الرتبة المئينية للمدارس الثانوية (HSPR) ودرجات اختبار الاستعداد الدراسي (SAT) ودرجات اختبار الشخصية للتنبؤ لاحقاً بمعدل النقاط الخاصة بالكلية. وأوضحت النتائج أن (HSPR) أفضل متنبئ منفرد، كما أن دمجاً موزوناً بشكل مناسب لكل من (HSPR) و (SAT) لمعدل النقاط الخاصة بالكلية أفضل تنبؤا من أي من المتغيرات المستقلة وحدها، وأن درجات اختبار الشخصية لا تقدم إضافة مفيدة للفائدة التنبؤية لمدمج (HAPR) و (SAT). وسوف يناقش الارتباط المتعدد بتفصيل أوسع فيما بعد في هذا الفصل.

وجهات نظر حول تفسير معامل ارتباط
Considerations for Interpreting A Correlation Coefficient

قد يكون حساب معامل الارتباط بسيطاً، لكن تفسيره قد يكون محيرا. فهو ربما يكون أحد أكثر الإحصاءات التي أسيء تفسيرها و / أو فسرت بشكل مفرط من الباحثين. وهناك آراء مختلفة تحتاج إلى أن تؤخذ بنظر الاعتبار عند تقييم حجم الارتباط وفائدته العملية.

الحجم Size

يجب أن يقرر الباحث ما إذا كان الارتباط "عاليا" بما يكفي لهدف بحثي معين. وبالوسع تقييم الحجم بطرق ثلاثة: 1) حسب حجمه المطلق وفائدته التنبؤية، 2) بالعلاقة مع الارتباطات الأخرى بالمتغيرات ذاتها أو المشابهة لها أو 3) في إطار دلالته الإحصائية.

الحجم المطلق والصدق التنبؤي: عملياً، إن التصنيفات الخاصة بوصف درجة أهمية العلاقة بين المتغيرات هي كما يلي:

قيمة (r)	العلاقة
.86 – 1.00	عالية جدا
.70 - .85	عالية
.50 - .69	معتدلة
.20 - .49	متدنية / منخفضة
.00 - .19	مهمل

إن أحد عيوب هذا التصنيف هو أن الحدود الفاصلية هي حدود تعسفية / فرضية. فمن الواضح أن (r) البالغ (.84) يشبه (r) البالغ (.86) أكثر مما يشبه (r) البالغ (.75) رغم أن (.84) هو في صنف الـ (.70) ذاته.

إن مؤشر تقييم القوة النسبية للعلاقة الذي لا ينطوي على أصناف تعسفية هو "معامل التحديد (r^2)". (أنظر الفصل 5). ويوضح مربع معامل الارتباط هذا نسبة التباين التي يشترك فيها المتغيران (X) و (Y). فإذا، مثلاً، كان (r) بين درجات اختبار الاستعداد للرياضيات ودرجات الرياضيات (.50) فإن نسبة التباين في درجات الحساب التي تتسبب فيها درجات اختبار الاستعداد هي ($R^2 = .50^2 = .25$). أما بقية الـــ 75% للتباين فهي تمثل النسبة التي لا يفسرها الارتباط. ويوضح لنا معامل التحديد مدى إمكانية تنبؤنا بمتغير واحد عن طريق استخدام درجات متغير آخر. فإذا كان الارتباط بين درجات الجبر والاختبار (أ) هو (.40) والاختبار (ب) هو (.60) والاختبار (ج) هو (.80) فإن تباين الدرجات المشترك مع الاختبار (أ) هو (.16) والاختبار (ب) هو (.36) ومع الاختبار (ج) هو (.64). لاحظ أننا لو أخذنا

بنظر الاعتبار الطبيعة الهندسية / النمطية لمعامل التحديد فإن الفرق في فائدة التنبؤ لـ (r = .6) و (r = .8) هو أكبر من الفرق في الفائدة التنبؤية لـ (r = .6) و (r = .4).

مقارنة مع معاملات ارتباطات أخرى: يجب تقييم حجم الارتباط بالنسبة لمعاملات ارتباط أخرى للمتغيرات ذاتها أو لمتغيرات مماثلة. مثلاً، (r) البالغ (.75) سيعتبر متدنياً بالنسبة للعلاقة بين نتائج شكلين متكافئين لاختبار تحصل، لكنه يعتبر عالياً بالنسبة للعلاقة بين مقياس الاستعداد ودرجات طلبة الكلية الجدد. وعلى مدى السنوات فإن كلاً من اختبار الاستعداد الدراسي (SAT) واختبارات الكليات الأمريكية (ACT) قد ارتبطت بصورة نموذجية بمعدلات طلبة الكلية الجدد بحوالي (r = .40). ويقع الارتباط (.40) ضمن صنف متدن في تصنيفنا العملي. ومع ذلك، فإن 16% من التباين في معدلات نقاط طلبة الكلية الجدد الذي يمكن التنبؤ به من أي من هذين الاختبارين هو تباين مفيد بما يكفي لاتخاذ قرارات قبول بما يسوّغ استخدامها. إن أي شخص بمقدوره أن يبني اختباراً للطلبة الأقدمين في الثانوية له معامل ارتباط (.50) مع معدل نقاط طلبة الكلية الجدد يمكن له أن يصنع ثروة من تسويقه لأنه سيعزى إليه 25% أي، ($.50^2$) من GPA للطلبة الجدد، مقابل 16% أي ($.40^2$) لكل من ACT و SAT. إن إحدى طرق تقييم حجم الارتباط تكون بتعريفه عالياً إذا كان أعلى مما ينافسه، ومتدن إذا كان أدنى مما ينافسه. وللمساعدة في تقييم الارتباطات الحاصلة، ينبغي على الباحث من خلال مراجعة الأدبيات تحديد أية مستويات من الارتباط توجد نموذجياً من أجل أنواع محددة من البيانات.

الدلالة الإحصائية: ينبغي لدى تقييم حجم الارتباط، الأخذ بعين الاعتبار حجم العينة التي يستند إليها الارتباط. فدون معرفة حجم العينة، لا يعرف المرء ما إذا كان الارتباط قد حدث نتيجة صدفة فحسب، أو إذا ما كان هناك احتمال علاقة أصيلة. فإن كان هناك اقل من 20 حالة في العينة (وهو أمر لا نوصي به) فإن (r) "متواضع" مثل (.50) يمكن أن يحدث بسهولة عن طريق الصدفة. وينبغي أن يكون المرء حذرا في إعطاء أهمية كبيرة للارتباطات الكبيرة عندما تكون حجوم العينات المشمولة صغيرة. ولأجل تجنب خطا استنتاج علاقة غير موجودة حقا، فإنه ينبغي على الباحث أن يصوغ الفرضية الصفرية بأن ارتباط المجتمع الإحصائي يساوي (0) ثم يحدد ما إذا كان الارتباط الناتج يبتعد بشكل كاف عن (0) لتسويغ رفض الفرضية الصفرية. وقد تعرضت في الفصل (6) إلى كيفية استخدام الجدول (A.5) في الملحق، الذي يدرج قيماً حرجة للمعامل (r) لعدد مختلف من درجات الحرية (d f). وبمقارنة الناتج (r) مع القيم الحرجة لـ (r) المدرجة في الجدول، يمكن للمرء أن يحدد الدلالة الإحصائية للارتباط التتابعي. مثلاً، افترض أن دراسة ارتباطيه تشمل (92) تلميذا تمخضت عن ارتباط مقداره (.66). القيمة الحرجة المدرجة في الجدول هي (.27) على مستوى (.01) و (.21) على مستوى (.05). ولان الارتباط الحاصل (.66) يفوق القيم المجدولة فإن بوسع الباحث أن يذكر أن الارتباط دال إحصائيا. فالفرضية الصفرية سترفض والباحث سيستنتج أن متغيرين مرتبطان في المجتمع الإحصائي.

يمكن أن نرى من الجدول (A.5) أنه كلما كانت العينة صغيرة، لابد أن يكون الحجم المطلق للارتباط كبيرا كي يبلغ الدلالة الإحصائية. ويبين الجدول (A.5) أنه مع (22) حالة (df=20) ، هناك حاجة لمعامل ارتباط (.54) كي يكون دالاً على مستوى (.01). ومع (102) حالة، على أية حال، فإن ارتباط (.25) دال على مستوى (.01) وعندما N=1000 فإن (r) المساوية (.08) يعد دالاً على مستوى (.01).

الفائدة العملية: يجب على المرء أن يدرس الأهمية العملية لمعامل الارتباط. فرغم أن معامل الارتباط دال إحصائيا فقد تكون له منفعة عملية قليلة. فلو كان لدى باحث عينة من (1000) فإن معاملاً صغيراً جدا مثل (.10) سيكون دالاً إحصائيا على مستوى (.01). لكن، أي أهمية عملية سينطوي عليها هذا الارتباط؟ المعلومات حول (X) فقط تعزوه إلى (1%) من التباين (r^2) في Y. وفي هذه الحالة قلما يكون من الأهمية، ذلك الإرهاق، بجمع درجات (X) للتنبؤ حول (Y). أما مدى أهمية الارتباط فإنها تعتمد على فائدته التنبؤية بالنسبة لكلفة الحصول على بيانات التنبؤ. فالتنبؤ ذو الارتباط العالي الصعب والباهظ للحصول عليه قد ينطوي على قيمة عملية أقل من تنبؤ رخيص وسهل، ذي ارتباط متدن.

وينبغي الملاحظة أن معامل الارتباط يصف، فقط، درجة العلاقة بين تعريفات إجرائية لعامل التنبؤ والمتغيرات المتنبأ بها في وضع بحثي خاص لعينة محددة من الأفراد. ويمكن بسهولة أن تتغير قيمته إن قيست المتغيرات نفسها وترابطت باستخدام تعريفات إجرائية مختلفة و / أو عينة مختلفة.

الارتباط والعلّية CORRELATION AND CAUSATION

يتجلى في تقييم الدراسات الارتباطية، أحد أكثر الأخطاء الشائعة، وهو تفسير الارتباط وكأنه يدل بالضرورة على علاقة علّية – علة ونتيجة. ويشير (Saslow; 1982) إلى هذه الممارسة من تفسير الارتباط كعلاقة علية، كواحد من "الخطايا السبع المهلكة" التي يرتكبها الباحثون. مثلاً، إذا وجد ارتباط موجب دال بين عدد ساعات مشاهدة التلفاز وأوزان تلامذة المدارس المتوسطة، فإن ذلك لا يبرهن بأن الإفراط في مشاهدة التلفزيون يسبب السمنة. تذكر من الفصل (10) انه عندما لا يكون المتغير المستقل تحت سيطرة الباحث، فإنه ينبغي عندئذ النظر في تفسيرات بديلة. وفي هذا المثال تبدو العلية العكسية معقولة. فربما، كلما ازداد وزن الطفل، مال إلى اختيار مشاهدة التلفاز بدلاً من النشاطات البدنية والألعاب والتفاعل مع الأقران. ويبدو تفسير العلة المشتركة معقولاً أيضاً. وربما تعود الاختلافات في أنماط الترفية العائلية وأسلوب الحياة، إلى الفروقات في كل من الوزن والوقت المصروف على مشاهدة التلفاز.

وينبغي التأكيد، على أية حال، أنه ليس صحيحاً بأنه لا يمكن، مطلقاً، للارتباط، تقديم

دليل يربط بين علة ونتيجة. فتحذير رئيس دائرة الصحة حول مخاطر التدخين يستند إلى دراسات وجدت ارتباطات إيجابية بين عدد السجائر المدخنة كل يوم والإصابة بسرطان الرئة والأمراض الأخرى. وهنا، كما ذكرنا في الفصل (10) فإن العلّة العكسية (السرطان يؤدي إلى تدخين السجائر) لا يعد تفسيراً موثوقاً. ويتضح أن فرضيات العلّة المشتركة (مثل: الناس الذين يعيشون في مناطق التلوث الجوي العالي يدخنون اكثر ولديهم معدلات إصابة أعلى بسرطان الرئة) لا تشكل مثل هذه الحالة. ولان الدراسات الارتباطية مع البشر تتفق مع نتائج الدراسات التجريبية مع الحيوانات فإن تحذير رئيس الدائرة الطبية يعتبر ذا أساس قوي.

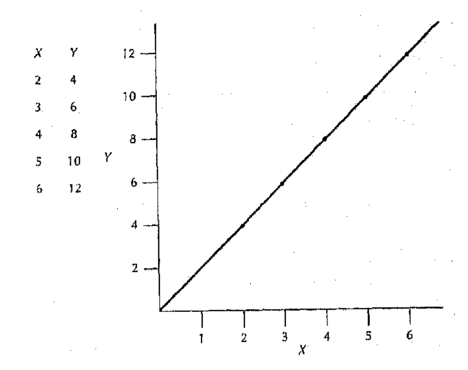
رياضيات التنبؤ MATHEMATICS OF PREDICTION

إذا وجد أن متغيرين مرتبطان، فإن من الممكن استخدام أحد المتغيرين للتنبؤ بالآخر. وكلما كان الارتباط أعلى كان التنبؤ أدق. مثلاً أوضحت البحوث الارتباطية أن الدرجات العالية في الثانوية ومقاييس الاستعداد الدراسي ذات صلة / ترتبط بمعدل النقاط GPA في الكلية. فإذا أحرز طالب درجة عالية في اختبارات الاستعداد وكانت له درجات عالية في الثانوية فمن المحتمل أن يحرز في الكلية درجات أعلى من الطالب الذي يحرز درجات متدنية في متغيري التنبؤ. وبوسعنا التنبؤ وبدرجة دقة معينة (معدل النقاط) GPA المحتمل من درجات الثانوية ودرجات اختبار الاستعداد. ولا ينطبق هذا التنبؤ في كل حالة لأن العوامل الأخرى كالدافعية، أو المبادرة، أو عادات الدراسة، لا تؤخذ بنظر الاعتبار. ولكن بشكل عام، يعتبر التنبؤ دقيقاً بما يكفي لأن يكون مفيداً.

دعنا نوضح التنبؤ بجملة بسيطة من البيانات. فالشكل 11.1 يبين رسماً بيانياً لبيانات مترابطة. ففي مثل هذه الرسوم البيانية يتم تحديد عامل التنبؤ (X) دائماً على المحور الأفقي، بينما المتغير المعيار(*) / التابع "criterion" (Y) على المحور الرأسي. (عامل التنبؤ "predictor" هو متغير يُستمد منه التنبؤ والمعيار هو متغير يتم التنبؤ به). المتغيران (X) و (Y) في هذا المثال يترابطان بشكل تام ($r = 1.00$) وهكذا تقع جميع النقاط على الخط القطري المعروف بـ "خط الارتداد". وبوسعنا استخدام الخط فيما بعد للتنبؤ من قيم (X) المعروفة بقيم (Y). مثلاً، يمكن لأحدهم أن يرى أن الدرجة (12) في المتغير (Y) يتم التنبؤ بها بالنسبة لأشخاص لديهم درجة (X) تساوي (6).

(*) المتغير المعيار (Criterion variable): هو المتغير الذي يكون "مثار الاهتمام الرئيس" للبحث. فالباحث يهدف إلى فهم ووصف هذا المتغير والتنبؤ به من خلال تأثير متغيرات أخرى (عوامل تنبؤ) فيه - وتسميته الشائعة هي "المتغير التابع" وللأسف، فإن كلمة معيار (criterion) التي تطلق أيضاً اختصاراً على "المتغير التابع" تشيع في البحوث بما قد يربك القارئ إذا لم يدرك المفهوم الخاص للكلمة في السياق - (المراجع).

الشكل (11.1): بيانات افتراضية تبين ارتباطاً تاماً



وبدلاً من رسم مخطط بياني للتنبؤ فإن ما هو مناسب أكثر تطوير معادلة للتنبؤ. وهذه المعادلة هي حالة خاصة للمعادلة الرياضية للخط المستقيم وهي كما يلي:

$$Y' = bX + a \qquad (11.2)$$

حيث:

Y'= درجة التنبؤ في Y

b = ميل خط الارتداد (مقدار التغير في Y لكل وحدة تغير في X)

a = مقطع Y(*) (حيث يقطع الخط محور Y، وهو قيمة درجة Y عندما X=0)

في المثال أعلاه، يمكن أن نرى أن (b) (الميل) هو (2). ولكل تغيير وحدة في (x) يتغير (Y) بمقدار (2).

$$b= \frac{Y2 - Y1}{X2 - X1}$$

عندما يتغير (X) من (3) إلى (4) يتغير (Y) من (6) إلى (8):

$$b= \frac{8-6}{4-3} = 2$$

(*) مقطع Y (intercept) مصطلح يشير إلى قيمة Y عندما X=0

ويمكن تحديد مقطع Y (a) بفحص الرسم البياني. فالخط يقطع محور (Y) في (0). وبالتعويض في المعادلة Y'=bX +a يمكن للمرء أن يتنبأ بقيمة (Y) المناظمة لكل قيمة في (X). فإذا كانت (X) تساوي (5) فإن درجة (Y) التنبأ بها هي Y'=2(5) + 0 = 10.

في المثال أعلاه يرتبط (X) و (Y) بشكل تام وكان من السهل تحديد (b) و (a). لكن في اغلب أوضاع التنبؤ لا يترابط (X) و (Y) بصورة تامة. دعنا ندرس مثالاً بسيطاً آخر في الشكل 11.2. يمكننا أن نرى في الشكل 11.2A أنه لا تقع كل النقاط على خط مستقيم. إلا أنه يمكن رسم خط ارتداد من خلال النقاط يمثل أفضل خط توفيقي لتلك البيانات. لكنه ينبغي أن نطور أولاً معادلة الارتداد. فقيم (a) و (b) في المعادلة تتحدد حسب معيار "المربعات الصغرى". ويؤدي هذا المعيار إلى خط ارتداد يرسم لكي يكون متوسط مربع الانحرافات لدرجات (Y) الفعلية من الدرجات المتنبأ بها على الخط أقل ما يمكن. وعندما نستخدم معامل ارتباط بيرسون (r) لتنبؤ فإن ميل خط الارتداد (b) يساوي معامل ارتباط بيرسون مضروباً بالانحراف المعياري للدرجات المتنبأ بها (σ_y) مقسومة على الانحراف المعياري لدرجات التنبؤ (σ_x):

$$b = r\frac{\sigma_y}{\sigma_x} \qquad (11.3)$$

وفي المثال أعلاه:

$$b = .48\left(\frac{1.85}{1.41}\right) = 1.10$$

ومقطع y (a) هو وسط الدرجات المتنبأ بها ($\overline{Y}$) ناقصاً ناتج ضرب الميل (b) في وسط عامل التنبؤ ($\overline{X}$):

$$a = \overline{Y} - b\overline{X} \qquad (11.4)$$

وفي المثال أعلاه:

$$a = 5.6 - 1.10(3) = 2.3$$

والآن يمكننا أن نحسب الدرجة المتنبأ بها ($\overline{Y}$) لكل قيمة لـ (X) بضرب الميل (b) في درجة (X) وإضافة مقطع Y وهو (a):

$$Y' = bX + a \qquad (11.5)$$

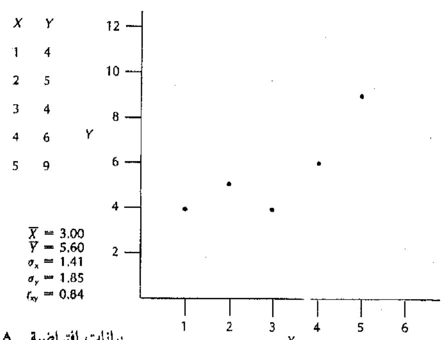
وفي المثال أعلاه،

$$Y' = 1.10X + 2.3$$

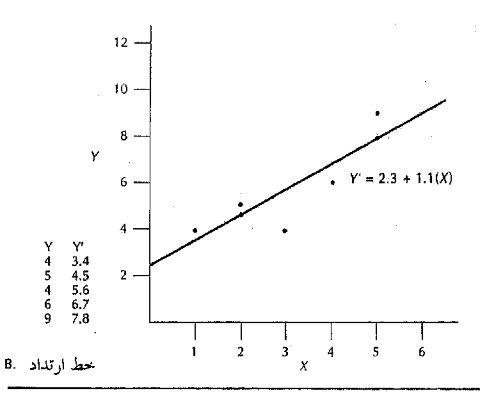
Y'=2.3+1.1(5)=7.8 Y'=2.3 + 1.1(2)=4.5

ونظراً لأن كل درجات Y' المتنبأ بها تقع على خط الارتداد فيمكننا من قيمتي Y' المساويتين (4.5) و (7.8) أن نرسم خط ارتداد كما هو مبين في الشكل (11.2B).

الشكل (11.2): بيان تشتت بيانات افتراضية مع خط ارتداد



A. بيانات افتراضية



B. خط ارتداد

تذكر أن خط الارتداد يمثل قيم (Y) المتنبأ بها حسب كل قيمة لـ (X). ويمكن للمرء أن يرى في الشكل (11.2B) أن درجات Y الفعلية لا تقع على الخط. ويمثل التباين بين درجات (Y) الفعلية ونقاط ('Y) المتنبأ بها حسب قيمة (X) المعينة تمثل خطا التقدير (e):'e = Y - Y. وباستخدام خط الارتداد يمكن أن يتنبأ المرء أن Y=4.5 عندما X=2. لكنه يمكننا أن نرى أن درجة (Y) الفعلية كانت (5). ومن ثم فإن خطأ التقدير هو:
(0.5 = 0.5 =4.5 - e=5). ان مجموع الأخطاء هو دائما (0). وأن التباينات السالبة توازن بالضبط التباينات الموجبة.

لحساب مؤشر متوسط الخطأ، كما في حساب الانحراف المعياري، يربع المرء كل (e) ويجمع (Σe^2) ثم يقسمها على (n - 2). وبأخذ الجذر التربيعي لهذا العدد، يحصل المرء على انحراف معياري لأخطاء التقدير. ويفيد هذا المؤشر الذي يدعى "الخطأ المعياري للتقدير" كمؤشر على دقة التنبؤ. ويمكن كتابة المعادلة كما يلي:

$$\sigma_{est} = \sqrt{\frac{\Sigma(Y - Y')^2}{n-2}} \qquad (11.6)$$

حيث:

σ_{est} = الخطا المعياري للتقدير

Y = قيم Y الفعلية حسب قيمة X المحددة.

'Y = قيم Y المتنبأ بها حسب قيمة X المحددة.

n = عدد الملاحظات.

وهناك معادلة بديلة اسهل حسابيا في استخدامها لحساب خطأ التقدير المعياري:

$$\sigma_{est} = \sigma_Y \sqrt{1 - r^2} \qquad (11.6a)$$

حيث:

σ_Y = الانحراف المعياري لدرجات (Y).

r^2 = مربع معامل الارتباط بين (X) و (Y).

لننظر في مثال ينطوي على التنبؤ. لاحظت مشرفة على مرشدي مدارس ثانوية أن بعض موظفيها ممن تنقصهم الثقة بأنفسهم بدوا أنهم يؤدون أعمالهم بشكل اضعف ممن لم يكونوا كذلك. هكذا قررت أن تطور مقياسا للثقة بالنفس لترى ما إذا كان ذلك المقياس مفيداً للتنبؤ حول نجاح المرشدين.

لقد حددت عناصر متنوعة للمفهوم البنائي "عدم الثقة بالنفس"، ومع نصيحة ومساعدة مرشدين محترفين آخرين صنعت مقياسا لعدم الثقة بالنفس ذي صدق بناء وثبات مقنعين. ثم

طبقت المقياس على عينة تمثل مرشدي المدارس الثانوية. وقد سجلت لكل واحد من الأفراد الوسط الحسابي لرضى الزبون على مقياس توضح فيه الأرقام العالية رضى أكبر.

وجدت المشرفة معامل بيرسون (-.50) بين درجات مقياس عدم الثقة بالنفس ورضى الزبون. ودلّها اتجاه الارتباط انه كلما ازدادت درجات عدم الثقة مالت درجات الرضى إلى النقصان. وأوضح الفرق بين الارتباط و (0) مدى قوة العلاقة. وتم إيضاح الفائدة التنبؤية بمعامل التحديد (r^2). في هذه الحالة كان (r^2) هو $(.25 = .5^2-)$ موضحاً أن (25 %) من التباين في درجات الرضى كانت مشتركة مع درجات عدم الثقة بالنفس. إن بوسع المشرفة التنبؤ بدرجات رضى الزبون (Y') باستخدام درجات المرشدين لعدم الثقة بالنفس (X). الخطوة الأولى هي حساب (b) في معادلة التنبؤ. وكان لمقياس عدم الثقة بالنفس وسط حسابي يبلغ (20) وانحراف معياري هو (4) بينما كان لدرجات الرضى وسط (16) وانحراف معياري (3). وهكذا

$$b = r\left(\frac{\sigma_x}{\sigma_y}\right) = -.5\left(\frac{3}{4}\right) = -.375$$

الخطوة التالية هي حساب (a)

$$a = \overline{Y} - b\overline{X}$$
$$= 16 - (-.375)20$$
$$= 23.5$$

فمعادلة التنبؤ بدرجات (Y) من درجات (X) هي هكذا:

$$Y' = bX + a$$
$$= -.375X + 23.5$$

فالمرشد الذي له الدرجة 24 على مقياس عدم الثقة بالنفس سوف نتنبأ له درجة رضى تبلغ 14.5 = 23.5 + 24 (-.375). أما بالنسبة لدرجة 14 على مقياس عدم الثقة فسوف يتنبأ المرء بأنها تساوي[*] 18.25=23.5+14(-.375). إن فحصاً مفيداً لدقة حساب (b) و (a) يكون حساب درجة (Y) المتنبأ بها بالنسبة للوسط في (X). وسوف ينتج عن الحسابات الصحيحة وسط (Y). وفي هذه الحالة 16 = 23.5 + 20(-.375).

(*) كل التنبؤات دقيقة عندما نتعامل مع الأفراد الذين حصلنا على بياناتهم. وعند استخدام هذه الإجراءات للقيام بتنبؤات حول أفراد تتوفر X لهم فقط، فإن فائدة هذه الإجراءات تتحدد بصدق الافتراض بأن الارتباط للأفراد الجدد يظل مماثلاً للارتباط الأصلي. فمثلاً، إذا استخدمت المشرفة درجات اختبار عدم الثقة بالنفس لتقرير من سنختار من بين المتقدمين لوظيفة الإرشاد، فإن درجات هذا الاختبار مفيدة إذا ما كان الافتراض بأن الارتباط لدرجات هذا الاختبار بين المتقدمين مماثلاً إلى (-.5) الذي وجد في الدراسة الأصلية.

الخطأ المعياري للتقدير Standard Error of Estimate

بما أن الارتباط في مثالنا ليس تاماً، فإن بعض المرشدين سيحرزون درجات في رضى الزبون أعلى مما هو متوقع، إذا أخذنا بنظر الاعتبار درجات مقياس عدم الثقة بالنفس، كما أن بعض المرشدين سيحرزون درجات أدنى مما هو متوقع. إن الخطأ المعياري للتقدير يوضح المقدار الذي يتوقع فيه أن تختلف فيه درجات (Y) الفعلية عن درجات (Y) المتنبأ بها.

إن الخطأ المعياري للتقدير هو جوهرياً مقياس التباين، ومن ثم يمكن تفسيره كانحراف معياري. فحوالي 68% من درجات (Y) الفعلية يتوقع لها أن تقع بين زائد أو ناقص خطأ معياري واحد للتقدير (2.60)، وحوالي 95% يتوقع لها أن تقع بين زائد أو ناقص خطأين معياريين (5.20) من الدرجة المتنبأ بها. مثلاً، إن أي مرشد له الدرجة (10) في مقياس عدم الثقة بالنفس، يمكن التنبؤ بأن تكون له درجة رضى تبلغ (19.75). وباستخدام الخطأ المعياري للتقدير يتوقع المرء أن (68%) من هؤلاء المرشدين الذين نتنبأ لهم (19.75) ستكون لديهم درجات رضى بين (17.15) و (22.35). و نتوقع أن (95%) سيكون لديهم درجات ما بين ($\pm 2\sigma_{est}$) أو بين (14.55) و (24.95).

إجراءات الارتباط متعدد المتغيرات
MULTIVARIATE CORRELATIONAL PROCEDURES

الارتداد المتعدد Multiple Regression

في العديد من أوضاع الحياة الحقيقية، يستخدم أكثر من متغير للتنبؤ بمعيار/ بمتغير تابع(*) معين. ويدعى التنبؤ بمعيار معين، باستخدام متغيري تنبؤ أو أكثر "الارتداد المتعدد". مثلاً، يمكن التنبؤ بمعدل نقاط الكلية، بصورة نموذجية، عن طريق استخدام جميعة / توليفة متغيرات يقدمها الطلبة في عملية القبول. أما معادلة الارتداد المتعدد فهي امتداد لمعادلة الارتداد البسيط:

$$Y' = a + b_1X_1 + b_2X_2 + \ldots b_nX_n \qquad (11.7)$$

حيث:

Y' = المعيار / المتغير التابع المراد التنبؤ به

a = ثابت

$b_1, \ldots b_n$ = وزن/ قيمة الارتداد لكل عامل تنبؤ

X_1 إلى X_n = درجة لكل عامل تنبؤ.

(*) المعيار/ المتغير التابع (Criterion): هذا مصطلح هنا يشير إلى "المتغير قيد الاهتمام" في دراسة الارتباط وهو المتغير الذي يتجه اهتمام التنبؤ نحوه كما أشرنا في هامش سابق – (المراجع)

يروز (يتفحص بعناية) الإجراء الإحصائي كل عامل تنبؤ لكي تعطى متغيرات التنبؤ في الجميعة أفضل تنبؤ للمتغير التابع ('Y) . إن استخدم الارتداد المتعدد موضح في دراسة (Jacobs, 1985).

لقد استخدم جاكوبس بيانات (4145) طالباً جديداً في جامعة بالوسط الغربي للتنبؤ بمعدل نقاط الفصل الأول. وكانت قيم عامل التنبؤ درجات في اختبارات الاستعداد الدراسي الفرعية (SAT لفظية و SAT كمية) إضافة إلى رتبة المدرسة الثانوية النسبية للطلبة (RHSR). فالرتبة النسبية للمدرسة الثانوية يتم الحصول عليها بتقسيم مرتبة الطالب في صف التخرج في الثانوية على حجم الصف، وتعدل حسب التغير في حجم الصفوف المتخرجة. ويبين الجدول (11.5) الارتباطات البسيطة بين كل من عوامل التنبؤ والمعيار/ المتغير التابع. ويمكننا أن نرى من الجدول (11.5) أنه ليس لأي من المتغيرات ارتباط عال مع (GPA) للطلبة الجدد، وأن عامل التنبؤ الأفضل هو الرتبة النسبية للثانوية. وقد تم تطبيق تحليل الارتداد المتعدد بغية تحديد أفضل ارتباط ممكن للجميعة المرجحة لمتغيرات التنبؤ الثلاثة مع (G PA). لقد كان الارتباط المتعدد (R) هو (.50). وهكذا فإن 25% $\left(R^2\right)$ من التباين في (GPA) كان يعزى إلى جملة درجات (SATV) و (SATH) و (RHSR). فكان للجميعة/ التوليفة المرجحة من المتغيرات أعلى ارتباط مع GPA وهكذا قدمت أدق تنبؤ. وكانت معادلة الارتداد لهذه البيانات هي:

$$Y'\,(GPA) = 2.813 + (-.0131)RHSR + .0014SATM + .0008SATV$$

وتوضح المعادلة أن (GPA) المتنبأ به لكل فرد يساوي مجموع الثابت (2.0813) مع نواتج ضرب كل من عوامل التنبؤ الثلاثة في وزنها الارتدادي. افترض أن طالباً كان لديه الدرجات التالية: SATV = 460 SATM = 540, RHSR = 21. أن قيمة GPA المتنبأ به للطالب في الجامعة سيكون (2.93):

$$Y' = 2.0813 + (-.0131)(21) + .0014(540) + .0008(460) = 2.93$$

وكان الخطأ المعياري للتقدير هو (.55). وبالنسبة لهذا الطالب كان الاحتمال (.68) بما يعني أن معدل نقاطه الفعلي سيكون بين (2.38) و (3.48).

الجدول 11.5 ارتباطات كل عامل تنبؤ مع المعيار / المتغير التابع

RHSR(*)	SATM	SATV	
- 0.42	0.39	0.31	GPA

(*) يعود الارتباط السالب بين الرتبة النسبية لطالب المدرسة الثانوية ومعدل النقاط (GPA)، إلى طريقة قياس الرتبة في الصف. فرتبة الطالب الأعلى تحصيلاً "1" (أقل الأعداد) ورتبة الطالب الأدنى تحصيلاً في الصف تعادل حجم الصف (عدد أكبر). فيتوقع أن يكون للطلاب الذين رتبتهم العددية متدنية، أعلى معدل نقاط (GPA)، ومن ثم كان معامل الارتباط سالباً.

يجب أن يكون لكل حالة في تطور الارتداد المتعدد أو معادلة التنبؤ، درجة لكل متغير من المتغيرات المستقلة (X) (عوامل التنبؤ) والمتغير التابع (Y) (المعيار). وكذلك يمكن في المستقبل استخدام المعادلة مع مجموعات مماثلة حين تعرف المتغيرات المستقلة (X_s). ويمكن استخدام درجات (SAT) ورتب المدرسة الثانوية النسبية للتنبؤ بقيمة GPA المجهولة للطلبة الجدد المتقدمين للجامعة.

ينبغي قياس المتغيرات على سُلّم / مقياس فترات. ومن الممكن على أية حال أن نضع متغيرات نوعية مثل الجنس، والطبقة الاجتماعية، والحالة الزوجية، والاتجاه السياسي، وما شابه في معادلة ارتداد إن تم ترميزها إحصائياً كمتغيرات ثنائية. مثلاً، إذا كان المتغير هو الجنس (Gender)، فإن الأعداد (1) يمكن تخصيصها للإناث والأصفار للذكور. إن مثل هذه المتغيرات التي جرى ترميزها إحصائياً يشار إليها في الارتداد المتعدد كمتغيرات صورية / وهمية.

ونظراً لأن الحسابات معقدة جداً، فإن الارتداد المتعدد يتم بالحاسوب. فبرامج الحاسوب متيسرة بحيث لا توفر معامل الارتباط المتعدد (R) ومعادلة الارتداد فحسب، بل كذلك نسبة التباين في المعيار/ المتغير التابع عن جميعة عوامل التنبؤ (R^2)، واختبارات الدلالة الإحصائية للمعامل (R) ومساهمة كل عامل تنبؤ.

ثمة بضع إجراءات متيسرة لاختيار المتغيرات المستقلة في معادلة الارتداد المتعدد. الطريقة الأكثر شيوعاً في الاستخدام تدعى "الاختيار التدريجي". إن أول متغير يؤخذ بنظر الاعتبار للدخول هو المتغير ذو الارتباط الأكبر، أكان موجباً أو سالباً، مع المعيار. ويستخدم الاختبار الفائي (F) للدلالة لتحديد ما إذا كان المتغير، في المجتمع الإحصائي، يساهم بصورة دالة في التنبؤ بالمعيار(*). فإذا أوفى المتغير الأول بالمعيار، فإن المتغير الذي يضيف القسم الأعظم للتنبؤ يختار تالياً. ويستند الاختيار إلى أعلى ارتباط بعد الضبط الإحصائي للارتباط بين عامل التنبؤ (1) والمعيار (الارتباط الجزئي). أما المتغير الثالث المراد إدخاله فهو المتغير ذو الارتباط الأعلى بعد ضبط عاملي التنبؤ الأولين إحصائيا. وتستمر خطوات اختيار المتغيرات إلى أن لا توفي أية متغيرات أخرى بمعايير الدخول. وفي كل خطوة يحسب (R) الذي يبين ما إذا كان المتغير الذي أدخل سيضيف على نحو دال إلى مقدار التباين في المعيار الذي تم التنبؤ به عن طريق المتغيرات التي تم إدخالها. وعلى الباحث أن يرجع إلى مرشد المستخدم الذي يصاحب برنامج/ رزمة الحاسوب الذي سيتم استخدامه (مثلاً $SPSS^x$) في وصف الطرق الأخرى لاختيار المتغيرات في الارتداد المتعدد.

(*) لتحديد ما إذا كان المتغير سيدخل، تجري مقارنة قيمة F حسب معيار مقرر سلفاً. فأحد المعايير الأول هو القيمة الصغرى للإحصاءة F التي ينبغي أن يحققها المتغير كي يكون دالاً. ويحدد المعيار الثاني الاحتمال المقترن بالإحصاءة F، مثل 0.05. ويدخل المتغير في المعادلة عندما يكون فقط، الاحتمال المقترن باختبار F أصغر من أو يساوي مستوى الدلالة المحدد.

وإذا شاء المرء استخدام معادلة الارتداد المتعدد لأهداف التنبؤ، فلابد من تحقيق صدقها في عينة أخرى. ونظراً لأن أخطاء الصدفة تعمل بصورة مختلفة في عينات مختلفة، فإن أوزان الارتداد المحسوبة للعينة الأصلية، قد لا تكون هي نفسها في عينة أخرى، أي أن جميعة / توليفة عوامل التنبؤ الأفضل في العينة الأصلية، قد لا تكون الأفضل في عينة أخرى. ولغرض تحديد صدق معادلة الارتداد المتعدد، فلابد أن تكون "صادقة بالتقاطع" عن طريق تطبيقها على عينة أخرى. ففي الصدق المتقاطع / التبادلي سيكون الارتباط المتعدد أدنى، عموماً، من العينة الأصلية التي حسبت فيها أوزان الارتداد. إن هذا الميل في الارتباطات المتعددة نحو النقصان عند تكرار البحث مع عينة مختلفة يشار إليه بالانكماش / التضاؤل. ويعتمد مقدار الانكماش على حجم العينة الأصلية وعدد عوامل التنبؤ (في الواقع، نسبة عوامل التنبؤ إلى حجم العينة). فكلما كان حجم العينة التي تستمد منها أوزان الارتداد، كبيراً كان الانكماش أقل. ولغرض الحصول على عوامل ارتباط متعددة (R_s) تكون مستقرة فإنه يوصى بشكل عام أن يكون لدى الباحث (30) فرداً على الأقل لكل عامل تنبؤ في معادلة الارتداد. ويوصي بعضهم أن يكون هناك (300) فرد لكل عامل تنبؤ.

لا يتقيد الارتداد المتعدد بأهداف التنبؤ. فهو شائع الاستخدام في أنواع مختلفة من البحوث. ونظراً لأن بمقدور الارتداد المتعدد أن يحدد المتغيرات التي تسبب التباين في المتغيرات التابعة، فإن بالإمكان استخدامه لمساعدة الباحثين في فهم الظواهر المعقدة. فيمكن للباحث الذي يود فهم ظواهر مثل النجاح المهني أو السياسة الليبرالية المحافظة استخدام الارتداد المتعدد لتحليل المساهمات المنفصلة والجمعية لعدد من المتغيرات المستقلة لهذه الظواهر.

التحليل المميِّز / التمييزي Discriminate Analysis

يعد التحليل المميز إجراءاً إحصائياً مرتبطاً بالارتداد. فهو يستخدم عدداً من متغيرات التنبؤ لتصنيف الأفراد إلى مجموعتين متميزتين أو أكثر، مثل المتسربين إزاء المواظفين على الدراسة أو الطلبة الناجحين إزاء غير الناجحين أو الجانحين إزاء غير الجانحين وما إلى ذلك. فالمعيار في التحليل المميز هو انتماء الشخص للمجموعة. ويؤدي هذا الإجراء إلى معادلة أو دالة مميزة حيث تضرب درجات عوامل التنبؤ بالأوزان بما يسمح بتصنيف الأفراد إلى مجموعات. وحين تكون هناك مجموعتان فقط فإن الدالة المميزة هي أساساً معادلة ارتداد متعدد مع معيار انتماء في مجموعة يرمز له بـــ (0) أو (1). غير أنه مع ثلاث مجموعات أو أكثر حسب المعيار، فإن التحليل المميز يتجاوز الارتداد المتعدد.

استخدم (Vacc & Picot, 1984) التحليل المميز لتحديد عوامل التنبؤ للنجاح في كلية ذات برنامج للدكتوراه في التربية. وأرادا إيجاد المتغيرات التي ميزت الانتماء في واحدة من اثنتين من المجموعات: الذين اكملوا بنجاح دراسة الدكتوراه والذين لم يكملوها. وقد استخدما عدداً

من عوامل التنبؤ مثل درجات اختبار ميلر للقياس MAT و GPA لطلبة الكليات و GPA للخريجين وفترة الوقت بين درجة الماجستير والالتحاق ببرنامج الدكتوراه، واختصاص الدكتوراه، والعمر لدى الالتحاق، والجنس، والحالة الزوجية، والعرق.

إن المتغيرات التي اعتبرت عوامل تنبؤ ناجحة في هذه الدراسة كانت درجة MAT والجنس Gender وحقل التخصص. وكانت المعادلة المميزة هي:

$$Z = C + .57382X_{MAT} - .44895X_{sex} + .65722X_{major} \qquad (11.8)$$

توضح الإشارة الموجبة لمعامل MAT أن الطلبة ذوي الدرجات العالية في MAT هم الأكثر احتمالاً في إكمال برنامج الدكتوراه. ونظراً لأن المؤلفين لم يوضحا كيف تم ترميز الجنس، فيستحيل من المعادلة تحديد أكبر الاحتمالين لكل من الذكور أو الإناث لإكمال الدراسة. وكان المعامل سالباً، ومن ثم فإن الجنس ذو الرقم الرمزي المتدني، يحتمل له أن يكون أكثر نجاحاً على الأرجح. وقد ذكر الباحثان أن الاختصاص في الإدارة التربوية كان الأكثر احتمالاً لإكمال الدراسية. ووجد أن المعادلة المشتقة أكثر فاعلية في تحديد الطلبة الناجحين (95.6% تم تحديدهم بشكل صحيح) من الطلبة غير الناجحين. وأعتقد الباحثان أن عدد الطلبة غير الناجحين كان ضئيلاً جداً بما لا يسمح التحليل في ضوئه.

إن الحسابات المعقدة في التحليل المميز تتطلب حاسوباً. ويعتبر تفسير النتائج معقداً كذلك، كما لابد للباحث المبتدئ من استشارة مختص بالإحصاء له معرفة بهذا المجال.

التحليل العاملي Factor Analysis

ثمة إجراء آخر شائع الاستخدام وقائم على الارتباط وهو التحليل العاملي. فهذا الإجراء يحلل الارتباطات المتداخلة بين مجموعة كبيرة من المقاييس بغية تحديد عدد صغير من "العوامل" المشتركة. والعوامل هي مفاهيم بنائية افتراضية حيث يجرى افتراضها لتشكل أساساً لأنماط المقاييس النفسية المختلفة كالذكاء، والاستعداد والتحصيل والشخصية، ومقاييس الاتجاهات. ويوضح التحليل العاملي المدى الذي تقيس فيه الاختبارات أو الأدوات الأخرى الشيء نفسه، وتمكن الباحثين من التعامل مع عدد صغير من القيود. إن بعض دراسات التحليل العاملي لاختبارات الذكاء، مثلاً، حددت عوامل تشكل أساساً لها: شفوية، وعددية، ومكانية، وذاكرية، واستدلالية.

وتنطوي الخطوات الأولى للتحليل العاملي على اختيار المتغيرات التي سيشملها التحليل وتطوير مصفوفة الارتباط التي تبين ارتباط كل مقياس مع كل مقياس آخر. وقد يكون هناك عدد كبير جداً من الارتباطات في المصفوفة. وتخضع المصفوفة إلى حسابات في برنامج حاسوبي للتحليل العاملي ينتج عنه عناقيد من المتغيرات التي تترابط فيما بينها بصورة عالية، لكن هناك

ارتباطات متدنية بين العناقيد. وهذه العناقيد هي العوامل، والهدف هو تحديد عدد صغير من عوامل منفصلة تقع تحتها ويمكنها أن تفسر التباين الاقتراني بين عدد كبير من المتغيرات. دعنا، مثلاً، ندرس مصفوفة الارتباط الافتراضية التالية:

	1	2	3	4	5	6
1- مفردات	–	.80	.15	.20	.22	.25
2- أقيسة (Analogies)	–	–	.12	.25	.10	.28
3- حساب	–	–	–	.75	.12	.12
4- استدلال عددي	–	–	–	–	.20	.22
5- إكمال الصورة	–	–	–	–	–	.82
6- تصميم الكتلة	–	–	–	–	–	–

يبين "تدقيق" في التحليل العاملي ثلاثة عناقيد. فالارتباط بين المفردات والأقيسة عال (.80) إلا أن أياً منهما لا يرتبط بصورة عالية مع المتغيرات الأخرى. ويرتبط الحساب والاستدلال العددي بصورة عالية مع بعضيهما، ولكن ليس مع المتغيرات الأخرى. ويمكن رؤية نمط مشابه مع إكمال الصورة وتصميم الكتلة. ويبدو أن هناك ثلاثة عوامل تحدد الأداء في هذه المقاييس الستة – شفوي وعددي ومكاني.

ثم يتم بعدئذ حساب الارتباطات بين كل من المتغيرات الأصلية والعوامل المشتركة التي تم تحديدها. وتدعى هذه الارتباطات "ارتباطات العامل"(*). ويوضح حجم ارتباطات العامل أهمية العامل على الأداء في كل متغير. ويأمل المرء أن يجد لكل متغير ارتباطاً عال مع عامل معين ارتباطات متدنية مع العوامل الأخرى. وسيشير لك إلى أن المتغير يشترك في التباين مع العوامل الأخرى في ذلك العامل، لكنه يشترك في تباين قليل أو معدوم مع المتغيرات المفروضة على العوامل الأخرى.

تنطوي الخطوة التالية على ما يدعى "تدوير العامل". وتعد تلك عملية تبسيط مصممة لإنتاج نمط واضح من العوامل والمتغيرات التي ترتبط بها. ونتيجة لذلك تكون العوامل أكثر تفسيراً بصورة مباشرة. وتعرض برامج الحاسوب طرق التدوير المختلفة ولكل واحدة منها معيار للتبسيط. ويجب على المستخدم أن يتخذ قراراً حول الطريقة التي يريد استخدامها. ثم يقوم الباحث بعدئذ بتسمية العوامل الناتجة. وتنطوي هذه الخطوة على تحديد المتغيرات التي ترتبط بشكل دال مع عامل معين ويستمد اسماً ينطبق على جميع المتغيرات. وقد يؤدي تحليل ما إلى عامل عددي، أو عامل لفظي، أو عامل مكاني...الخ.

(*) ارتباط العامل (Factor Loading) ويستخدم مصطلح (loading) بمعنى الارتباط، للإشارة إلى ارتباط عامل، مثل السمة الشخصية، بدرجة أداء مستمدة من اختبار نفسي – (المراجع).

لقد أخضع (Posner & Kouzes, 1988) أداة تدعى أداة جدول ممارسات القيادة (LPI) للتحليل العاملي. وقد قاما بتحليل استجابات (2876) فرداً للثلاثين بندا في الأداة، واستمدا خمسة عوامل قيست عن طريق البنود الثلاثين. ويوضح الجدول (11.6) بنية العامل لـ (LPI). ويمكن أن نرى أن للبنود (8,18,23,13,28,3) ارتباطات عالية مع العامل (1) وارتباطات متدنية مع العوامل الأخرى. وتبدو هذه البنود كأنها تقيس قدرة القائد على تمكين الآخرين على الفعل. ومن ناحية أخرى فإن البنود (5,25,15,20,10,30) لها ارتباطات عالية مع العامل (2) وهكذا.

وفي الواقع، فإن تغطية الإجراءات الحسابية لأي من التحليل المميز أو التحليل العاملي تقع خارج نطاق هذا الكتاب. وندعو القارئ إلى مراجعة كتب مثل "مقدمة في التحليل العاملي" لمؤلفيه (Kim & Muller 1978) أو كتاب "التحليل المميز" (Klecka, 1980) أو كتاب "التحليل العاملي" (Gorsuch, 1983) من أجل مناقشات يسيرة لهذه الموضوعات.

الارتباط الممتد Canonical Correlation

يعتبر الارتباط الممتد تعميماً للارتداد المتعدد الذي يضيف أكثر من متغير تابع إلى معادلة الارتداد المتعدد. إنك تتذكر بأن معامل الارتباط المتعدد يبين الارتباط بين "أفضل" توليفة/ جميعة من المتغيرات المستقلة ومتغير واحد تابع. ويمدّد، الارتباط الممتد، التحليل إلى أكثر من متغير واحد تابع. وبعبارة أخرى، فإن الارتباط الممتد تحليل ارتدادي مع عدة متغيرات مستقلة وعدة متغيرات تابعة. وهو يأخذ بنظر الاعتبار درجات (X) و (Y) والعلاقات بين المتغيرات (X)، وبين المتغيرات (Y)، وبين مجموعات المتغيرات (X) و (Y). وتكون النتيجة معامل ارتباط ممتد يمثل أقصى ارتباط ممكن بين مجموعات درجات (X) ومجموعات درجات (Y). ويوضح أيضاً المساهمات النسبية للمتغيرات المستقلة وغير المستقلة المنفصلة في الارتباط الممتد. لذا بوسع المرء أن يرى أي المتغيرات التي تكون الأكثر أهمية بالنسبة للعلاقات بين المجموعات.

ويجري الارتباط الممتد دائما في الحاسوب. فالحسابات اليدوية شاقة جداً وبالغة التعقيد. وللاستزادة بمعلومات أكثر حول الارتباط الممتد، يمكن العودة إلى تحليل الارتباط الممتد (Thompson, 1984).

الجدول 11.6 بنية العامل (ارتباطات عاملية) لأداة / استبانة ممارسات القيادة

(N = 2876)

البنـــد	العامل 1 تمكين الآخرين على الفعل	العامل 2 إنعاش الفؤاد	العامل 3 الهام رؤية مشتركة	العامل 4 تحدي الطريقة	العامل 5 نمذجة الطريق
8	.719	.173	.096	.008	.098
18	.694	.200	.176	.088	.214
23	.680	.198	.189	.231	.273
13	.526	.169	.092	.085	.006
28	.509	.280	.206	.195	.290
3	.459	.208	.235	.069	.256
5	.111	.731	.220	.099	.109
25	.152	.725	.255	.143	.128
15	.402	.689	.102	.129	.113
20	.451	.673	.163	.148	.172
10	.400	.635	.079	.154	.189
30	.224	.532	.194	.250	.240
7	.185	.215	.709	.251	.119
2	.156	.165	.657	.276	.136
27	.223	.255	.623	.384	.239
17	.173	.225	.615	.270	.240
22	.223	.151	.506	.362	.136
12	.166	.114	.481	.345	.107
16	.180	.169	.266	.641	.233
26	.164	.185	.241	.637	.057
11	.043	.082	.184	.622	.145
1	.182	.128	.219	.648	.153
21	.354	.194	.178	.473	.145
6	.170	.049	.138	.392	.173
29	.218	.185	.144	.192	.609
9	.343	.158	.031	.107	.512
14	.164	.164	.239	.228	.509
4	.232	.142	.353	.238	.411
19	.109	.156	.334	.315	.409
24	.319	.120	.115	.227	.372

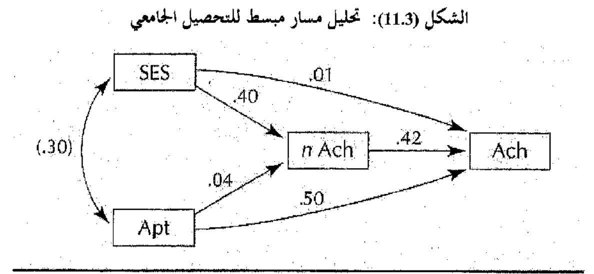
Source: From "Development and Validation of the Leadership Practices Inventory" by B.Z. Posner and J.M. Kouzes, 1988, *Educational and Psychological Measurement, 48*, p. 493.

تحليل المسار Path Analysis

إن تحليل المسار جملة من الإجراءات الإحصائية مصممة لاختبار نموذج علّي افتراضي حول العلاقات بين ثلاثة متغيرات أو أكثر. واستناداً إلى النظرية والمعرفة الراهنة يقترح الباحث نموذجاً علياً ثم يطبق تحليل المسار كي يقرر ما إذا كان النموذج العلّي متسقاً مع البيانات التجريبية. أما النماذج التي لا تتطابق/ تتسق مع البيانات فتُرفض، بينما ينظر إلى النماذج التي لا ترفض بكونها أنماطاً علّية ممكنة كي تكون عرضة لاستقصاء أكثر.

مثلاً، افترض أن باحثاً مهتماً بالتنبؤ حول التحصيل الجامعي، وأن لديه نظرية تنص على أن الدافعية تعد متغيراً مهما يؤثر على التحصيل، ومن ثم يجب إدخاله في كل دراسة تنبؤية(*). ويختار الباحث عينة ويحصل على قياسات دقيقة لجميع المتغيرات المعنية. وبعد حساب معاملات الارتباط بين جميع أزواج المتغيرات، يرسم الباحث مخطط مسار يوضح فيه السياق الذي يفترض أن تؤثر فيه المتغيرات على بعضها (النموذج العلّي). ويبين الشكل 11.3 التفسير المفترض للباحث حول التحصيل الجامعي (Ach) باستخدام الحاجة للتحصيل (*n* Ach) ومؤشر الدافعية إضافة إلى الاستعداد (Apt) والحالة الاجتماعية الاقتصادية (SES) كمتغيرات مستقلة. ويستخدم المخطط أسهماً تسمى "المسارات" لإيضاح اتجاه التأثير بين المتغيرات، ومن هنا جاء اسم تحليل المسار على هذا الإجراء.

الشكل (11.3): تحليل مسار مبسط للتحصيل الجامعي



يشير الخط المستقيم ذو الرأس السهمي الواحد إلى التأثير المباشر لمتغير على آخر - مثلاً، يؤثر (SES) على (*n* Ach) مباشرة. ويمثل أي تأثير غير مباشر بسهم يأتي من متغير هو بحد ذاته

(*) هذا المثال المبسّط لتحليل المسار مأخوذ من (Kerlinger, 1979).

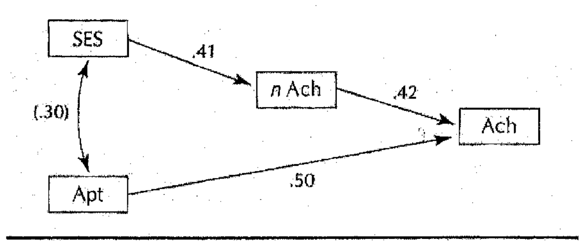
"متلق" لسهم آخر – مثلاً (SES) يؤثر على التحصيل بصورة غير مباشرة من خلال (n Ach). لذا يوضح الشكل 11.3 أن كلاً من (SES) والاستعداد يؤثران على دافعية (n Ach) مباشرة، وإضافة إلى ذلك يؤثر كلاهما على التحصيل بشكل غير مباشر، من خلال تأثيرهما المباشر على (n Ach). وتشير الخطوط المنحنية ذات الرؤوس السهمية في كلا الطرفين، إلى المتغيرات التي أقيم الارتباط بينها مع افتراض أن تغيرها يعود إلى علل خارج نطاق النموذج. ورغم أن (SES) و (Apt)، مثلاً، مترابطة (.30) فإنه لا توجد أي متغيرات في النموذج يُفترض أنها تؤثر على (SES) و (Apt)، ولا توجد هناك فرضيات حول الكيفية التي ترتبط فيها مع بعضها علّياً.

يقوم الباحث بعدئذ بإجراء تحليل الارتداد الذي يُنتج أوزان الارتداد التي تدعى "معاملات المسار". وتتضح هذه المعاملات على المسارات في الشكل 11.3. مثلاً، ينتج ارتداد (n Ach) (متغير تابع) حسب (SES) والاستعداد (متغيرات مستقلة)، معامل مسار مقداره (.40) بالنسبة للمسار من (SES) إلى (n Ach) و (.04) بالنسبة للمسار من الاستعداد إلى (n Ach). وتوضح هاتان القيمتان أن تأثير (SES) كبير على (n Ach) لكن تأثير الاستعداد ضئيل. أما المعامل من (SES) إلى التحصيل فهو (.01) بينما المعامل من الاستعداد إلى التحصيل فهو (.50) ومن (n Ach) إلى التحصيل (.42).

إن ما يسمح به تحليل المسار، بما يزيد على الارتداد المتعدد الاعتيادي، فهو إن تحليل المسار يبين التأثيرات غير المباشرة العاملة، وكذلك التأثيرات المباشرة. فالمتغيرات الثلاثة في التحليل تؤثر على التحصيل مباشرة، إلا أن تحليل المسار يبين كذلك أن (SES) والاستعداد يؤثران على التحصيل بشكل غير مباشر من خلال (n Ach) (الدافعية).

ولن يقف الباحث عند هذا الحد، بل سيختبر مخطط المسار في الشكل 11.3 ويحذف المسارات ذات المعاملات الزهيدة (من SES إلى Ach، .01) ومن (Apt حتى n Ach، .04). ويعاد بعدئذ حساب معاملات المسار كما هي مبينة في الشكل 11.4 وتجري مقارنتها بمعاملات التحليل الأول. ويتم طرح المصفوفة الثانية للارتباطات من مصفوفة الارتباطات الأصلية للحصول على الفروقات. فإذا كانت الفروقات صغيرة جداً استنتج الباحث أن النموذج الممثل في تحليل المسار النهائي أبسط وأكثر اقتصاداً، كما يقدم أفضل تفسير للظاهرة المراد التنبؤ بها. ويبين الجدول 11.7 مقارنة المصفوفات من هذه الدراسة. وتتضح الارتباطات الأصلية إلى اليسار، والارتباطات المحسوبة من معاملات المسار إلى اليمين، بينما الفروقات الحاصلة من طرح الاثنين في أسفل الجدول. ويمكن أن نرى أن مجموعات الارتباطات متشابهة، والفرق الأكبر هو 0.4 فقط. وهكذا فإن نموذج المسار في الشكل 11.4 متسق مع البيانات. ووفقا لتحليل المسار يستنتج الباحث أن الاستعداد يؤثر على التحصيل مباشرة، لكنه لا يؤثر على (nAch). وتؤثر (SES) على التحصيل فقط من خلال (n Ach).

الشكل 11.4 تحليل مسار مبسط للتحصيل الجامعي مع إعادة حساب معاملات المسار



الجدول 11.7 الارتباطات الأصلية، الارتباطات المعادة، ومصفوفة الفروق
(أربعة متغيرات، نموذج ثان)

الارتباطات الأصلية

	1	2	3	4
1	1.00	.30	.41	.33
2	.30	1.00	.16	.57
3	.41	.16	1.00	.50
4	.33	.57	.50	1.00

الارتباطات المعادة

	1	2	3	4
1	1.00	.30	.41	.32
2	.30	1.00	.12	.56
3	.41	.12	1.00	.48
4	.32	.56	.48	1.00

مصفوفة الفروق

	1	2	3	4
1	.00	.00	.00	.01
2	.00	.00	.04	.01
3	.00	.04	.00	.02
4	.01	.01	.02	.00

ويمكن أن ينجم عن تحليل المسار نتائج مضللة ما لم تكن هناك ظروف معينة. مثلاً، يجب أن تلبي البيانات افتراضات معينة (تقاس المتغيرات حسب سُلّم / مقياس فترة وعلاقات خطية بين المتغيرات، وما إلى ذلك)، ويجب قياس المتغيرات بأدنى الأخطاء، وأن يكون حجم العينة كافياً، كما يجب شمول المتغيرات العلّية المهمة في النموذج النظري.

يراد من هذا الوصف البسيط لتحليل المسار، جعل القارئ يألف الإجراء. فالحسابات

تتجاوز نطاق هذا الكتاب. ولأجل اطلاع أوسع على هذا الموضوع أنظر (Pedhazur, 1982) أو (Li, 1975) أو قم باستشارة إحصائي يعرف الإجراء.

أما بشأن وصف ميسر لدراسة منشورة تستخدم تحليل المسار، ينصح القارئ بالرجوع إلى (Culver, Wolfle & Cross, 1990). لقد اختبر هؤلاء الباحثون نموذج مسار عن العوامل المؤثرة على رضا المعلمين بعملهم في مراحل التحاقهم الأولى. فقد افترض النموذج أن الرضا بالعمل دالة للجنس، والسن، وثقافة الأب، وثقافة الأم، والتحصيل الأكاديمي للمدرس، وعدد سنوات التدريس، وتصوُّر المدرس للمدير وزملاء العمل، والتزام المدرس بمهنة التدريس. لقد أُجري التحليل بشكل منفصل على عينات من المدرسين السود والبيض. ولكلا المجموعتين، تم الاستنتاج بأن متغيرات الخلفية الديموغرافية كالعمر، والجنس، وثقافة الأب، وثقافة الأم كانت ذات أهمية ضئيلة للرضا بالعمل، مقارنة بمتغيرات مباشرة أكثر هي المناخ المدرسي والالتزام بالتدريس. وكان لسنوات الخبرة التدريسية تأثير قليل على مستوى الرضا بالعمل. وبصورة عامة كانت معاملات المسار متماثلة بالنسبة للمدرسين السود والبيض. وكانت الاستثناءات عندما ارتد الرضا على التحصيل الأكاديمي. فالمدرسون البيض ذوو الإنجاز المتدني كانوا يميلون إلى أن يكونوا أكثر اقتناعاً بأعمالهم من نظرائهم ذوي التحصيل العالي. أما بالنسبة للسود فلا توجد مثل هذه الفروقات. وثمة فرق آخر لوحظ حدوثه عند ارتداد الرضا على الجنس. كان الذكور السود يميلون إلى كونهم أكثر رضا بأعمالهم من الإناث السود، لكن الإناث البيض كن أكثر رضا من الذكور البيض. واستنتج الباحثون بان العملية المؤدية إلى الرضى بالعمل قد تكون مختلفة بالنسبة للسود والبيض. وقد أوصوا بإجراء بحوث أخرى حول هذه المسالة مع مجتمعات إحصائية ذات قاعدة أوسع من المدرسين.

الخلاصة SUMMARY

يتفحص البحث الارتباطي العلاقات القائمة بين المتغيرات. والإجراء الارتباطي شائع الاستخدام في البحوث التربوية و النفسية. فهو يمكننا من فهم أفضل لبعض الظواهر وإجراء التنبؤات. ويجب تفسير الارتباطات بشكل مناسب. فيجب على المرء أن ينظر في حجم العينة وحجم معامل الارتباط ودلالته الإحصائية والعملية. ولعل اخطر خطأ هو تفسير الارتباط بذاته كمؤشر على العلّية.

وتستخدم دراسات ارتباطية بسيطة متغيرين فقط لكل فرد في الدراسة. وقد ادرج هذا الفصل عدداً من الأنماط المختلفة لمعاملات الارتباط المستخدمة مع متغيرات تقاس حسب أنماط مختلفة من المقاييس. ويستخدم الارتداد المتعدد لإيجاد الارتباط بين متغيرين مستقلين أو أكثر ومتغير تابع. وتنشأ عنه معادلة ارتداد أو تنبؤ يمكن استخدامها فيما بعد للتنبؤ بالمتغير التابع بالنسبة لمجموعة جديدة من الأفراد، حينما تتوفر معلومات حول المتغيرات المستقلة.

لقد تعرض هذا الفصل بشكل موجز إلى عدد من الإجراءات الارتباطية الأكثر تعقيدا. ومن بينها الارتباط الجزئي، والتحليل المميز، والتحليل العاملي، والارتباط الممتد، وتحليل المسار. وينبغي أن يعرف القارئ نوع الوضع البحثي الذي تكون فيه كل من هذه الأساليب مفيدة.

مفاهيم أساسية Key Concepts

bacterial correlation	ارتباط ثنائي
canonical correlation	ارتباط ممتد
coefficient of correlation	معامل ارتباط
coefficient of determination	معامل تحديد
correlation matrix	مصفوفة ارتباط
correlational research	بحث ارتباطي
discriminate analysis	تحليل مميز / تمييزي
dummy variable	متغير صوري / وهمي
factor	عامل
factor analysis	تحليل عاملي
intercept	مقطع
multiple regression	ارتداد متعدد
partial correlation	ارتباط جزئي
path analysis	تحليل المسار
predictor	عامل تنبؤ
regression line	خط الارتداد
slope	ميل
standard error of estimate	خطأ معياري للتقدير

تمارين EXERCISES

1- ما الارتباط الذي يوضح أكبر مدى للعلاقة بين متغيرين؟

أ. 85. + أو 85.–

ب. 56. + أو 74.–

2- وجد موظف حكومي ارتباطا (.70) بين عدد الحانات في مدن الولاية وعدد الجرائم المرتكبة في هذه المدن. واستنتج أن تشريعا بتقليص عدد الحانات سيقلص من الجريمة. ماذا تقول عن استنتاج الموظف؟.

3- لماذا تعد تفسيرات معاملات الارتباط القائمة على مقاديرها فقط، مضللة أحيانا ؟

4- أرسم بيان تشتت للأزواج التالية من النقاط (X , Y).

X	2	3	10	6	4	9	7	3	6
Y	10	9	1	3	7	1	2	5	4

هل العلاقة:

أ. خطية أو خطية منحنية ؟

ب. عالية، معتدلة أو متدنية ؟

ج. موجبة أم سالبة ؟

5- أجرى باحث سلسلة من الاختبارات على مجموعة من الطلبة لدى التحاقهم بكلية القانون في الجامعة. وفي نهاية السنة الأولى من الكلية حصل الباحث على معدلات نقاط الطلبة.

أ. ما الأسلوب الارتباطي الذي ينبغي استخدامه لتحديد أقصى علاقة لهذه الاختبارات مع معدل النقاط GPA للسنة الأولى ؟

ب. ما الافتراضات التي ينبغي صياغتها بغية تعميم الاستنتاجات على المتقدمين الجدد لذات الكلية ؟ وعلى متقدمي كليات القانون الأخرى ؟

6- كيف تقارن ارتباطات (.90) و (.45) في إطار مقداريهما ؟

7- يهتم باحث بالقلق وكيف يؤثر على الأداء في اختبارات الاستعداد الدراسي.وكان لديه عالم نفس سريري يقيّم قلق الأفراد بترتيبه إياهم من (1) إلى (20). ثم يجرى اختبار استعداد مقنن لكل من الـ (20) فرداً ويحول درجات معامل الذكاء إلى رتب. ما معامل الارتباط الذي سيحسبه الباحث للبيانات؟ وضح إجابتك.

8- أثبت باحث ارتباطا مقداره (+.60) بين تقديرات مدراء مدارس لملابس المدرس والأداء الأكاديمي للطالب في (150) مدرسة ابتدائية في مقاطعته. واستنتج أن تشجيع المدرسين على ارتداء الملابس المناسبة سوف يزيد من الأداء الأكاديمي. علّق على هذا الاستنتاج.

9- حدد نوع مؤشر الارتباط المناسب للاستخدام في حل كل من المسائل التالية:

أ. ما هو معامل الارتباط بين الذكاء ودرجات الإبداع لمجموعة أفراد ؟

ب. ما هو معامل الارتباط بين أداء مجموعة أفراد في اختبارين حين يكون لدينا درجات

الأفراد في اختبار، ومستواهم حسبما يكون أعلى أو أدنى من المتوسط في الاختبار الآخر ؟

ج. ما هو الارتباط بين الجنس والتوظيف أو عدم التوظيف في عمل ما حين يكون لدينا (60) رجلاً و (60) امرأة ذوي درجات متطابقة في اختبار تأهيل لذلك العمل؟

د. ما هو الارتباط بين استجابات (180) تلميذاً في بنود اختبارين عندما نعرف الاستجابات الصحيحة والخاطئة لهؤلاء التلاميذ في بنود الاختبارين؟

هـ. ما هي العلاقة بين الاستعداد الدراسي ودافعية التحصيل عندما يكون لدينا درجات استعداد (200) طالب في اختبار ستانفورد - بينية ومستواهم في اختبار دافعية التحصيل باعتباره أعلى أو أدنى من المتوسط ؟

و. ما هي العلاقة بين الجنس والقدرة الميكانيكية حين يكون لدينا درجات (60) ولداً و (60) بنتاً في اختبار القدرة الميكانيكية ؟

ز. ما هي العلاقة بين درجات القياس الاجتماعي لمجموعة تلاميذ ومرتبتهم الأكاديمية في صف بمدرسة ثانوية ؟

10- إذا كنا سنحسب معامل ارتباط بين سمتين أ و ب، فأي مما يلي يجب أن يكون لدينا؟

أ. مجموعة واحدة من الأفراد يمتلك بعضهم صفات السمة أ وتمتلك البقية صفات السمة ب.

ب. قياسات السمة أ في مجموعة أفراد وقياسات السمة ب في مجموعة أخرى.

ج. مجموعة واحدة من الأفراد، بعضهم لديه السمتان أ و ب، وبعضهم ليس لديه أية سمة منهما، وبعضهم لديه سمة واحدة وليس الأخرى.

د. مجموعتان من الأفراد، واحدة يمكن تصنيفها أ أو ليس أ ، والأخرى ب أو ليس ب.

هـ. قياس السمتين أ. ب لكل فرد في مجموعة واحدة.

11- وجد بحث معين ارتباطاً موجباً بين درجات معامل الذكاء وطول فترة الانتباه بين أطفال بعمر العاشرة. ومن هذه النتائج، أي مما يلي سيكون استنتاجاً معقولاً ؟:

أ. فترة انتباه طويلة هي علة الذكاء.

ب. معدل ذكاء عال هو علة فترة انتباه طويلة.

ج. ثمة احتمال عال بأن عينة كبيرة من الأطفال بعمر العشر سنوات ذوي معامل ذكاء عال سيكون لديهم وسط فترة انتباه أقصر من عينة كبيرة من الأطفال بعمر العشر سنوات ذوي معامل ذكاء متدن.

د. قد يتنبأ المرء بفترات انتباه أطول بالنسبة لأطفال ذوي معامل ذكاء عال بعمر العشر سنوات من أقرانهم الأطفال ذوي معدل الذكاء المتدني.

12- اختبر موضوعات البحث التالية، وحدد ما إذا كان البحث التجريبي، أو العلّي - المقارن أو البحث الارتباطي هو التصميم المناسب لكل منها:

أ. تأثير طلاق الأبوين على دافعية الأداء لدى الأطفال.

ب. تأثير برنامج محدد لتدريس المفردات على تحصيل الدراسات الاجتماعية.

ج. العلاقة بين حجم الصف ورضا الطالب عن نوعية التدريس.

د. تأثير تدريس الصوتيات على المستوى/ العمر القرائي لطلبة الصف الرابع.

13- تود باحث ارتباطاً مقداره (-.42) بين الرتبة في صف بالمدرسة الثانوية والأداء في الكلية بالنسبة لعينة من (1500) طالب جديد. كيف تفسر هذا المعامل في إطار الاتجاه، والمقدار، والنسبة المئوية للتباين المشترك؟

14- تود موظفة إدارية في مدرسة التنبؤ بالتحصيل في اللغات الأجنبية. ويتوفر لها درجات في اختبار ذكاء، واختبار استعداد في اللغة، واختبار قراءة، كما أنها تعرف كذلك جنس (ذكر/ أنثى) الطلبة. كيف تنصحها في المتابعة للتنبؤ بتحصيل اللغات الأجنبية؟ وأي اسم سيعطى إلى هذا الإجراء؟.

15- أوضح كيف يمكن لقياسات متغيرين أن يرتبطا بطريقة منتظمة إلى حد ما دون وجود أية علاقة علّية بين المتغيرين؟.

16- يستقصي أحد الباحثين صدق البناء لاستبيان مصمم لقياس توتر المدرس. وهو يرغب في معرفة ما إذا كان الاستبيان يقيس مفهوماً بنائياً واحداً أو مفاهيم متعددة. ما الإجراء الذي توصي به الباحث ليجيب عن هذا السؤال ؟

17- أراد باحث التنبؤ بالتحصيل في السنة الأولى في كلية القانون. وقد أجرى تحليل ارتداد متعدد مع حجم عينة من (1000) واستخدم ستة متغيرات تنبؤ وهي: GPA لطلبة الكلية، وتخصص طلبة الكلية، درجة LAST، درجة اختبار الكتابة، الجنس، والوقت المنصرم بين درجة الكلية والتقدم إلى كلية القانون. وكانت نتيجة $R = .20$ دالة على مستوى دلالة (0.01). ما تفسيرك لنتائج الباحث؟.

18- افترض أن هناك ارتباطاً (+0.94) بين متوسط عدد الساعات المصروفة في الدراسة كل يوم (X) و (GPA) للفصل الأول للطلبة الجدد (Y) بالنسبة لمجموعة من (100) فرد. أحسب الـ (GPA) المتنبأ به للطالب الذي يدرس (6) ساعات في اليوم باستخدام المعلومات التالية:

$$\overline{X} = 5.74 \quad s_x = 2.31 \qquad \overline{Y} = 2.75 \quad s_y = 2.31$$

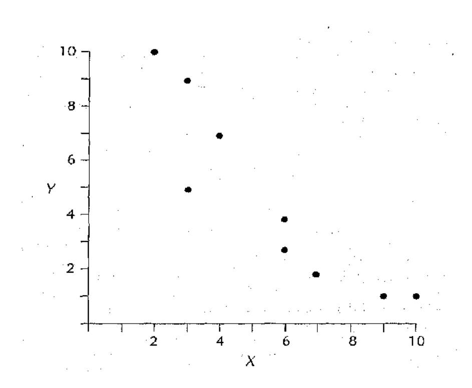
الأجوبة ANSWERS

1- أ. مدى العلاقة متماثل.
ب. 74. -

2- الارتباط بين متغيرين لا يوضح العلّية. في هذه الحالة، يحتمل أن تكون العلة المشتركة هي التفسير. فالمدن الكبيرة لديها حانات أكثر وجريمة أكثر، والمدن الصغيرة لديها حانات أقل وجرائم أقل.

3- لتفسير الارتباط، يحتاج الباحث إلى معرفة حجم العينة التي تستند عليها وحجم الارتباطات لنفس المتغيرات في دراسات أخرى.

4-



أ. خطي، ب. عالي، ج. سالب

5- أ. ارتداد متعدد.

ب. على الباحث افتراض أن العلاقة بين المتغيرات في المجتمع الإحصائي من المتقدمين إلى كلية القانون يبقى متماثلاً طوال الوقت. وعلى الباحث افتراض أن العلاقات بين المتغيرات بين المتقدمين لكليات القانون الأخرى متماثلة مع تلك العلاقات بين المتقدمين إلى كليتها.

6- قد يوصف r = .90 على أنه ارتباط عال جداً، و r = .45 متدن نوعاً ما. ولا يجب أن يقول المرء أن r = .90 هو أعلى مرتين من r = .45. وأفضل طريقة لمقارنتهما تكون في إطار معامل التحديد (r^2). ومع r = .90 فإن (.81) من التباين هو تباين مشترك. ومع r = .45 فإن (.20) فقط من التباين في (Y) يعود إلى التباين في (X). وهكذا فالفرق في حجم العلاقة هو (4) إلى (1) بدلاً من (2) إلى (1) كما قد يكون تم خطأ الافتراض من الحجم المطلق للمعاملات.

7- نظراً لان لدى الباحث بيانات ترتيبية أو نظام رتب فعليه أن يحسب معامل ارتباط سبيرمان للرتب.

8- ليس للباحث أي مسوغ لاستنتاج علاقة علّية على أساس الدليل الارتباطي فحسب. فتقديرات المدراء لملابس المدرس والأداء الأكاديمي للطالب يمكن أن يكونا تابعين لمتغير آخر ما.

9- أ. معامل ارتباط بيرسون (r) ب. ارتباط ثنائي ج. معامل فاي (Phi)
د. معامل فاي هـ. ارتباط ثنائي و. ارتباط ثنائي أصلي
ز. رو (rho) لسبيرمان

10- هـ

11- د

12- أ. بحث علّي مقارن ب. بحث تجريبي
ج. بحث ارتباطي د. بحث تجريبي

13- يوضح r = -.42 علاقة سالبة معتدلة بين المرتبة في صف المدرسة الثانوية والتحصيل في الكلية، أي أن الطلبة ذوي الترتيب العددي المتدني ينـــزعون نحو التحصيل العالي. ويعزى الارتباط السالب إلى الطريقة التي قيست بها المرتبة في الصف. إن لدى أعلى صاحب تحصيل في الصف المرتبة (1) (العدد الأدنى) ولدى أدنى صاحب تحصيل في الصف مرتبة مساوية لحجم الصف (أعلى عدد). فيتوقع للطالب ذي المرتبة (1) في صف من (400) أن يكون لديه (GPA) أعلى من الطالب ذي المرتبة (400).

وتدل (r = -.42) أن حوالي (18%) من التباين في درجات الكلية قد تعزى إلى التحصيل في المدرسة الثانوية كما تدل عليها المرتبات.

14- يختار الأداري في المدرسة عينة من الطلبة ويحصل على درجاتهم في الثلاثة اختبارات ودرجاتهم في اللغات الأجنبية. وسوف يرمز إلى الجنس بـ (0) أو (1). وسيتم إدخال البيانات في تحليل ارتداد متعدد مع اختبار الذكاء، واختبار الاستعداد اللغوي، واختبار

القراءة، والجنس كمتغيرات مستقلة، والدرجات في اللغة الأجنبية كمتغير تابع. ولدى تطوير معادلة الارتداد المتعدد فانه يمكن استخدامها للتنبؤ بتحصيل اللغة الأجنبية لمجموعات متشابهة من الطلبة عند معرفة المتغيرات المستقلة فحسب.

15- حتى عندما يتم ارتباط متغيرين لا يستطيع المرء الاستنتاج أن أحدهم يسبب الآخر. فالارتباط لا يدل على العلّية. وقد يتسبب متغير ثالث، لم يقم الباحث بقياسه، في العلاقة المرصودة.

16- يجب أن يخضع الباحث الاستبيان للتحليل العاملي.

17- مع حجم عينة كبير، لا يستغرب المرء أن معاملاً مقداره 0.20 فحسب، سيكون دالاً إحصائياً. وهي ليست كبيرة بما يكفي لتكون مفيدة في التنبؤ. ($R^2 = .04$) أي أن (4%) فقط من التباين في تحصيل كلية القانون سيكون قابلاً للتنبؤ من جميعة متغيرات التنبؤ هذه.

18-

$$b = .37$$

$$a = .63$$

$$Y' = a + bX$$

$$Y' = .63 + .37(6)$$

$Y' = 2.85$ (GPA المتنبأ به)

REFERENCES المصادر

Bourke, S. (1986). How smaller is better: Some relationships between class size, teaching practices, and student achievement. *American Educational Research Journal,* 23,558-571.

Culver, S.M., Wolfle, L.E., and Cross, L.H. (1990). Testing a model of teacher

satisfaction for blacks and whites. *American Educational Research Journal,* 27, 323-349.

Gorsuch, R.L. (1983). *Factor analysis.* Hillsdale, N.J.: Lawrence Erlbaum. Jacobs, L.C. (1985). *GPA prediction procedures and normative data for freshmen.* Bloomington: Indiana University, Bureau of Evaluative Studies and Testing.

Kerlinger, F.N. (1979). *Behavioral research: A conceptual approach.* New York: Holt, Rinehart and Winston.

Kim, J., and Mueller, C.W.(1978). *Introduction to factor analysis: What it is and how to do it.* Beverly Hills, CA: Sage Publications.

Klecka, W.R. (1980). *Discriminate analysis.* Beverly Hills, CA: Sage Publications.

Li, C.C. (1975). *Path analysis: A primer.* Pacific Grove, CA: Boxwood Press. Pedhazur, E.J. (1982). *Multiple regression in behavioral research: Explanation and prediction.* New York: Holt, Rinehart and Winston.

Posner, B.Z., and Kouzes, J.M. (1988). Development and validation of the Leadership Practices Inventory. *Educational and Psychological Measurement,* 48, 483-496.

Saslow, C.A. (1982). *Basic research methods.* Reading, MA: Addison-Wesley. Thompson, B. (1984). *Canonical correlation analysis: Uses and interpretation.* Beverly Hills, CA: Sage Publications.

Vacc, N.N., and Picot, R. (1984). Predicting success in doctoral study. *College Student Journal,* 18, 113-116.



البحث المسحي
Survey Research

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون بوسع الطالب أن:

1- يذكر هدف البحث المسحي.
2- يصف أربعة أصناف للمسوحات المصنفة حسب مداها وتركيزها.
3- يميز بين المسوحات الطولية والعرضية.
4- يصف الأنواع المختلفة للمسوحات الطولية.
5- يدرج الخطوات المعنية بتنفيذ المسح.
6- يوضح أهمية المعاينة الاحتمالية في البحث المسحي.
7- يستخدم بيانات العينة لتقدير فترة الثقة حول معْلم / معْلمة (Parameter) المجتمع الإحصائي.
8- يوضح مزايا المقابلة كأسلوب لجمع البيانات.
9- يكتب أسئلة مفتوحة النهاية ومغلقة النهاية.
10- يدرج خمسة إرشادات لإجراء مقابلة.
11- يذكر مزايا ومساوئ الاستبيان كأسلوب لجمع البيانات.
12- يذكر عشر قواعد لكتابة بنود الاستبيان.

13- يدرج إرشادات للاتباع ذات صلة بصيغة الاستبيان.

14- يوضح مزايا استبيان الاختبار الميداني.

15- يوضح إجراءات المتابعة التي ينبغي استخدامها في المسح البريدي.

16- يكتب رسالة التفسير / الغلاف للاستبيان.

17- يوجز إجراءات لمعالجة غير المستجيبين بعد استخدام إجراءات المتابعة.

18- يناقش إجراءات تقييم صدق وثبات الاستبيانات و / أو المقابلات.

19- يوضح تحليلات البيانات المناسبة لبيانات المسح.

20- يوضح الإحصائيات المستخدمة في جدولة البيانات.

المسح، أسلوب بحثي تجمع فيه البيانات عن طريق طرح أسئلة على مجموعة أفراد يُدعون المستجيبين. وهي طريقة بحثية شائعة الاستخدام في علم الاجتماع، والأعمال، والعلوم السياسية، والحكومة، والتربية. إن مدى الموضوعات التي تغطيها المسوحات والأساليب المستخدمة قد ازداد بشكل كبير في السنوات القليلة الماضية. فقلما يمضي أسبوع لا نتعرض فيه إلى أدوات الأعلام حيث تنشر نتائج مسح ما. فاستفتاء غالوب الشهير، مثلاً، يقوم بمسح الرأي العام وقضايا أخرى مختلفة. كما يسأل المختصون بالسوق عن أي المنتجات التي نشتريها أو قد نشتريها، بينما يسأل أصحاب الاستفتاءات السياسية حول من يمكن أن نصوت له، كما تريد شبكات التلفزيون أن تعرف أي العروض التي نشاهدها.

تعتبر المسوحات مهمة أيضاً في الدراسات العليا. فالعديد من الجامعات لها معاهد مختصة بالبحوث المسحية مثل معهد جامعة مشيغان للبحوث الاجتماعية ومعهد بحوث الدراسات العليا في لوس أنجلوس UCLA. ويقوم الأخير بإجراء مسوحات عن طلبة الكلية الجدد بغية الحصول على المعلومات الخاصة بأسباب اختيار الطلبة لكلية معينة، والتخصص المتوقع، والمهنة وخطط التخرج، وآرائهم السياسية، واتجاههم الديني، والعديد من المظاهر الأخرى في حياتهم. في خريف عام 1993، مثلاً، قام معهد UCLA بإجراء مسح على (220,757) طالب جديد دخلوا إلى (427) معهدا في الولايات المتحدة. ولعل أحد الاستنتاجات المهمة من ذلك المسح، إضافة إلى مضامين الدراسات العليا، هو رقم قياسي للنساء اللاتي يخططن لمتابعة وإكمال الدراسة للحصول على شهادات عليا (Cage, 1994). والعديد من أطروحات الدكتوراه والكثير من البحوث المنشورة في المجلات التربوية تنطوي على طرق مسحية.

وتجري المقاطعات التعليمية العامة مسوحات لجمع البيانات مثل متوسط عبء المدرس، وعدد أطفال ما قبل المدرسة في المجتمع، وعدد الطلبة الذين يشاركون في النشاطات اللاصفية، وبيانات عن موضوعات كثيرة أخرى.

أنواع المسوحات TYPES OF SURVEYS

قبل المبادرة بالبحث المسحي يجب على المرء تحديد الصيغة المناسبة للاستقصاء المقترح. وتصنف المسوحات وفق تركيزها ومداها (مسوحات العينة والإحصاء) أو وفق الإطار الزمني لجمع البيانات (مسوحات طولية أو عرضية). إن الإحاطة بالخيارات ستمكن الباحث من اختيار الطريقة التي ستؤمن البيانات الأكثر فائدة.

المسوحات المصنفة وفق التركيز والمدى
Surveys Classified According To Focus And Scope

المسح الذي يغطي مجتمعاً إحصائيا برمته يدعى "إحصاء / تعداد السكان"، ومثال عليه هو الإحصاء الأمريكي الذي تقوم به الحكومة كل عشر سنوات. وفي البحوث، على أية حال، لا يشير المجتمع الإحصائي إلى جميع سكان القطر. ويستخدم "المجتمع الإحصائي" للإشارة إلى مجموعة أشخاص بكاملها تطبق عليهم استنتاجات الدراسة. ويحدد الباحث المجتمع المعني المحدد. ويصعب على الباحثين في الغالب، بل يستحيل عليهم دراسة مجتمعات إحصائية كبيرة جداً. وعليه يقومون باختيار جزء صغير أو عينة من مجتمع إحصائي للدراسة. فالمسح الذي يدرس جزءا من المجتمع الإحصائي يدعى "مسح العينة".

قد تتحدد المسوحات بجداول بسيطة من "المحسوسات / الملموسات" مثل ما نسبة الأطفال الذين يستقلون باصات المدرسة؟ وما هو متوسط الانضمام إلى الصف؟ إن نوع المسوحات الأكثر تحدياً، هو الذي يحاول قياس غير "اللامحسوسات - مفاهيم بنائية"، مثل الاتجاهات، أو الآراء، أو القيم، أو العوامل النفسية والاجتماعية الأخرى. وفي مثل هذه الدراسة، يجب على المرء أن يضع في البال، ليس فقط المهارات المنطوية في المعاينة المناسبة، بل كذلك المهارات المعنية بتحديد أو بناء المقاييس المناسبة واستخدام الدرجات في مثل هذه المقاييس لصياغة عبارات ذات معنى حول المفاهيم البنائية المعنية. وإذا قمنا بتصنيف المسوحات على أساس مداها (الإحصاء مقابل العينة) وتركيزها (المحسوسات مقابل اللامحسوسات) لأصبح لدينا أربعة أصناف:

1. إحصاء المحسوسات، 2. إحصاء اللامحسوسات، 3. مسح عينة المحسوسات، 4. مسح عينة اللامحسوسات.

ولكل نوع مساهماته ومشكلاته المتأصلة.

إحصاء المحسوسات A Census of Tangibles

عندما يبحث المرء عن معلومات حول مجتمع إحصائي صغير، كمدرسة واحدة، وعندما تكون المتغيرات المعنية محسوسة، يكون هناك تحد قليل في إيجاد الإجابات المطلوبة. فإذا أراد مدير مدرسة أن يعرف كم منضدة في المدرسة، وكم طفلاً يستقل باص المدرسة، أو كم مدرساً

يحمل شهادة الماجستير، فإن حساباً بسيطاً سوف يؤمن له المعلومات. ونظراً لأن الدراسة تشمل كل المجتمع الإحصائي، فإن بوسع المدير أن تكون له الثقة الكاملة بتطبيق الاستقراء التام. فالمتغيرات معرّفة جيداً وواضحة لإجراء قياسها، وكلما كان الحساب دقيقاً ونزيهاً، فإن بوسع المدير أن يقول دون خوف من التناقض "في الأول من أيلول، كان هناك (647) منضدة في مدرستنا" أو "65% من المدرسين لديهم شهادات ماجستير". وتكمن قوة الإحصاء لهذا النوع في عدم قابليته للدحض. ويكمن ضعفه في حصره بمجتمع واحد محدود في نقطة زمنية محددة. والمعلومات المتوفرة عن طريق هذا النوع من الإحصاء تكون ذات أهمية مباشرة لمجموعة محددة، إلا أن مثل هذه المسوحات تضيف القليل إلى مجمل المعرفة في مجال التربية.

إحصاء اللامحسوسات A Census of Intangibles

افترض أن مدير مدرسة يبحث الآن عن معلومات حول تحصيل الطلبة أو تطلعاتهم أو معنويات الأستاذ أو مواقف الوالدين نحو المدرسة. ستكون المهمة اصعب لأن الإحصاء يتعامل مع مفاهيم بنائية لا ترصد بشكل مباشر، بل يجب استنتاجها من قياسات غير مباشرة. تنشر الدراسة القومية للتقييم المدرسي (NSSE) أداة رأي مصممة لقياس مواقف وأراء الطلبة أو المدرسين والوالدين حول المدارس. إن إدارة هذه الأداة على جميع الطلبة أو المدرسين أو الوالدين في النظام المدرسي سيشمل إحصاءاً لغير المحسوسات.

وكمثال آخر لمثل هذا النوع من الإحصاء، هو برنامج اختبار التحصيل الذي تقوم به معظم المدارس. فيجري اختبار جميع الأطفال وتستخدم درجات الاختبار لمقارنة أدائهم مع المعايير / المتوسطات القومية وأدائهم السابق وما إلى ذلك. ويجب أن يكون المدير على معرفة حول طبيعة أدوات القياس المستخدمة وملاءمتها لقياس تحصيل الطلبة في المدرسة، ومن ثم لابد أن تسأل عن مدى ثبات الاختبارات، وما إذا كانت تقيس المفهوم البنائي نفسه للتحصيل الذي حددته أهداف المدرسة ومدى جودة قياسها لذلك المفهوم البنائي.

إن قيمة إحصاء اللامحسوسات، هو سؤال يدور، إلى حد بعيد، حول المدى الذي تذهب إليه أدوات القياس المستخدمة في قياسها الفعلي للمفاهيم البنائية المعنية. فالأدوات الجيدة، إلى حد ما، متيسرة لقياس الاستعداد والتحصيل في مجالات أكاديمية مختلفة. ويظل العديد من المتغيرات الأخرى صعب القياس. ونظراً لأننا نفتقر إلى الأدوات التي تستطيع قياس المفاهيم البنائية على نحو مفيد، فإن العديد من الأسئلة المهمة في التربية لم يجر التعامل معها بصورة ناجحة. فمتغيرات مثل نجاح المدرس، ودافعية الطالب، والتكيف النفسي، والقيادة ظلت صعبة التعريف والقياس بصورة إجرائية.

مسح عينة المحسوسات / الملموسات A Sample Survey of Tangibles

عندما يحاول الباحثون البحث عن المعلومات حول مجموعات كبيرة، فإن كلفة إجراء إحصاء

هي في الغالب صادّة / مانعة. وعليه تستخدم أساليب المعاينة وتستخدم المعلومات التي تم جمعها من العينة لإقامة استنتاجات حول المجتمع الإحصائي برمته. ولقد رأينا انه عندما تتم المعاينة بشكل جيد، فإن الاستنتاجات الخاصة بالمجتمع الإحصائي يمكن أن تكون موثوقة.

وكمثال معروف عن مسح عينة محسوسات هو تقرير (Coleman, 1966). فقد أجريت هذه الدراسة استجابة للقسم (402) من قانون الحقوق المدنية لعام (1964)، التي وجهت مسؤول التربية إلى إجراء مسح حول التفاوت في الفرص التربوية بين مختلف المجموعات في الولايات المتحدة. واشتمل مسح العينة على أكثر من (600,000) طفل في الصف الأول، والثالث، والسادس، والتاسع، والثاني عشر لقرابة (4000) مدرسة. وقد اعتبرت المدارس ممثلة، عموماً، لجميع المدارس الأمريكية الحكومية رغم أن هناك بعض التمثيل المقصود المبالغ فيه للمدارس ذات الطلبة القادمين من الأقليات.

من البيانات المتولدة من المسح، تم الاستنتاج أن 65% من السود التحقوا بالمدارس التي كان فيها (90%) من الطلبة السود، و (80%) من البيض التحقوا بمدارس تضم أكثر من (90%) من البيض. وعند إجراء المقارنات بخصوص حجم الصف، والتسهيلات المادية، ومؤهلات المدرسين، كان هناك فرق ضئيل بين المدارس التي تخدم مجموعات أجناس وعروق مختلفة. ومع ذلك، فإن هذه المتغيرات اختلفت بين المناطق الحضرية والريفية وبين المناطق الجغرافية. أما المتضررون من هذه المتغيرات فيبدو أنهم أطفال الريف والموجودون في الجنوب بغض النظر عن العرق.

مسح العينة لغير المحسوسات / الملموسات A Sample Survey of Intangibles

تعتبر استفتاءات الرأي العام أمثلة عن دراسات تقيس مفاهيم بنائية غير ملموسة. ولا يلاحظ الرأي بشكل مباشر، بل يجب الاستدلال عليه من استجابات الأفراد حسب استبيانات أو مقابلات. وبدأ استفتاء الرأي في ثلاثينات القرن الماضي وازداد بشكل كبير. فحيثما كان المستجيبون راغبين في الكشف عن تفضيلاتهم بصورة حرة قبل الانتخابات، مثلاً، كان مستطلعو الرأي دقيقين في استنباط الرأي العام الذي تنبؤوا به حول نتائج الانتخابات التالية. لقد وفرت هذه الاستفتاءات أمثلة ممتازة عن فائدة إحصائيات العينة في تقدير معْلمات / بارامترات المجتمع الإحصائي. وعلى أية حال، إن أحجم الذين يساندون مرشحاً ما عن كشف تفضيلهم، بينما الذين يساندون المرشح الآخر يشعرون بالحرية في قول ذلك، فإن خطأ كبيراً سيدخل في نتائج الاستفتاء. فمثلاً، الناس أكثر رغبة في القول انهم سيصوتون ضد شخص يحتل منصباً ما. فقبل انتخابات الرئاسة عام (1948) أظهرت استفتاءات عدة أن (ديوي) سيسبق صاحب المنصب، وهو (ترومان) غير أن العديد من الناس قالوا أنهم لم يقرروا بعد. وفاز صاحب المنصب وهو (ترومان) بالانتخابات. ومن الواضح أن اغلب الذين كانوا قد أشاروا إلى أنهم لم يقرروا حقاً، قد صوتوا لصالح ترومان. كما أن المستجيبين يجمعون عن كشف خيارهم الذي قد يبدو قائماً على المصلحة الذاتية، والانحياز، أو الافتقار إلى المعرفة حول القضايا.

أما كيف سيدلي بصوته شخص ما في الانتخابات، فهذا أمر غير ملموس، لكن ما يلاحظ في الاستفتاء يعد شيئاً ملموساً. وتقوم خدمات الشبكات الأخبارية التلفزيونية بشكل جيد في التنبؤ بالكيفية التي ستنتخب بها الولايات، عندما يقوم بذلك عدد قليل من الدوائر الانتخابية، وذلك لأنها قادرة على استخدام مقاييس ملموسة للعينة (أي الكيفية التي أشرت فيها بعض الاستفتاءات) للتنبؤ بانتخاب السكان. وعليه فالمخاطر هي فقط تلك التي تنطوي على تقدير معْلمات السكان من إحصاءات العينة(*). على أية حال، إن أمام مستطلعي الرأي الذين يقدرون كيف سينتخب السكان على أساس الكيفية التي يقول بها الناس انهم سيصوتون، عائق إضافي في قياس ما هو غير ملموس في وقت إجراء القياسات. إن مسوحات غير الملموسات محددة بحقيقة أن البيانات التي نجمعها تقيس بشكل غير مباشر المتغيرات التي نهتم بها. وتعتمد جدية هذا التحديد على مدى الجودة الذي تقيس فيه الملاحظات المتغير غير الملموس.

قد يدرس المسح ذاته الملموسات وغير الملموسات في الوقت ذاته. لقد طلب مؤلفو تقرير كولمان من الطلبة أن يجيبوا عن الاستبيانات وأجروا اختبارات ذكاء وتحصيل لأجل الوصول إلى استنتاجات حول الطبقة الاجتماعية، والقدرة، والتحصيل، وكذلك علاقة هذه المتغيرات مع بعضها ومع المتغيرات الملموسة في الدراسة.

المسوحات المصنفة وفق البعد الزمني
Surveys Classified According to The Time Dimension

هناك نوعان من المسوحات المصنفة طبقاً لزمن جمع البيانات: المسوحات الطولية التي تدرس التغيرات عبر الزمن، والمسوحات العرضية التي تركز على نقطة واحدة من الزمن.

المسوحات الطولية Longitudinal Surveys

تجمع المسوحات الطولية المعلومات في نقاط مختلفة من الزمن بغية دراسة المتغيرات خلال فترات متسعة من الزمن. مثلاً، قد يختار الباحث الذي يدرس تطور الاستنتاج الكمي لدى أطفال المدارس الابتدائية عينة من طلبة الصف الأول ويطبق مقياساً للاستنتاج الكمي. وتتم متابعة هذه المجموعة في مراحل متعاقبة من الصفوف ويجري اختبارها كل سنة لتقدير الكيفية التي تتطور بها قدرات الاستنتاج الكمي عبر الزمن. وتستخدم ثلاثة تصميمات مختلفة في بحث المسح الطولي وهي: دراسات المجموعة ذاتها، دراسات النـــزعة ودراسات الجماعة.

دراسات المجموعة ذاتها (Panel Study): في هذا النوع من الدراسة، يتم مسح الأفراد "أنفسهم" في أوقات مختلفة على فترة موسعة. ونظراً لأنه تجرى دراسة الأفراد أنفسهم

(*) إحصاءة العينة (Statistic): تقابل معْلم / معْلمة (Parameter) المجتمع الذي يفترض أن العينة تمثله حقيقة – (المراجع)

على مدار زمن معين، فبوسع الباحثين أن يروا التغيرات في سلوك الأفراد ويبحثوا أسباب التغيرات. ومثال هذا النوع، هو المسح حول التدخين الذي جري في جامعة انديانا وتابع المجموعة ذاتها من الأفراد منذ عام 1980 (Zuhl, 1994, Feb. 4). وقد قامت الدراسة الأصلية بمسح (8000) طالب ثانوية حول آرائهم بالتدخين. وسألتهم أيضاً عن آرائهم حول الأمور الأكاديمية والأصدقاء وإجراء اللقاءات. ووجد الباحثون أن الطلبة الأقل جدية حول العمل المدرسي والأكثر تمرداً كانوا الأكثر احتمالا في البدء بالتدخين من نظرائهم ذوي التوجه الأكاديمي. ومال المدخنون كذلك إلى أن يكون لهم أصدقاء يدخنون أيضاً. وفي 1987 و 1988، آجري الباحثون أول مسح متابعة لهم فاتصلوا بـــ (6200) من المجموعة الأصلية. لقد وجدوا أن الاتجاهات نحو التدخين قد تغيرت مما يعكس عبور الأفراد نحو حالة البلوغ. وابتدأ الباحثون بمسح آخر للأفراد ذاقم عام 1993، والذي يتوقع أن يستغرق إكماله 2-3 سنوات. وسوف يسألون مرة أخرى الأفراد عن رأيهم بالتدخين وسوف يستكشفون وقد اصبح العديد منهم أباء وأمهات، كيف يؤثر إنجاب الأطفال على تدخينهم وان كان هناك انتقال في التدخين بين الأجيال.

دراسة النزعة / الاتجاه (Trend Study): يتم في هذا النوع مسح أناس "مختلفين" من المجتمع الإحصائي نفسه في أوقات مختلفة. مثلاً، الباحثون الذين درسوا الاتجاهات القومية في تحصيل الرياضيات يقومون بمعاينة طلبة المدارس المتوسطة في فترات مختلفة ويقيسون أداءهم في الرياضيات. ورغم أنه لا يتم اختبار الأشخاص أنفسهم في كل مرة، إذا ما اختيرت عينات من مجتمع طلبة المدارس المتوسطة بشكل عشوائي، فإنه بالإمكان اعتبار النتائج كل مرة ممثلة لمجتمع المدارس المتوسطة الذي سحبت منه عينات الطلبة وتتم مقارنة الدرجات من سنة إلى أخرى لرؤية ما إذا كان هناك اتجاه / نزعة بشكل واضح. مثال آخر عن دراسة النـــزعة / الاتجاه هو المسح حول تناول المخدرات بين المراهقين الأمريكان الذي يجريه سنوياً معهد جامعة مشيغان للبحوث الاجتماعية. وقد كان باحثو مشيغان يجرون مسحاً على طلبة الثانوية في سنة التخرج سنوياً منذ عام 1975، وعلى طلبة الصف الثامن والعاشر منذ عام 1991. ولقد وجد احدث مسح جرى على (51,000) طالب في (400) مدرسة في البلد، ولأول مرة في أكثر من عقد من الزمن كانت هناك زيادة بدلاً من النقصان في استخدام مختلف المخدرات المحرمة (Blonston, 1994, Feb. 1). ولعل الهم الأكبر هو التغير في الاتجاهات نحو المخدرات. وأوضح المسح أن قليلاً من الطلبة لا يوافقون على تناول المخدرات، وأقل من ذلك يروها تشكل خطراً محدقاً.

دراسة الجماعة (Cohort Study): في هذا النوع من الدراسة، تتم ملاحقة مجتمع إحصائي محدد على طول فترة من الزمن. فبينما تقوم دراسة النزعة بمعاينة مجتمع إحصائي عام يتغير في عضويته عبر الزمن، تعاين دراسة الجماعة مجتمعاً محدداً لا يتغير أعضاؤه خلال فترة المسح. مثلاً، قد يتابع نظام مدرسي صفوف التخرج في مدرسة ثانوية عام 1993 طول الفترة

ويطرح عليهم أسئلة حول الدراسة العليا، وخبرات العمل، واتجاهاتهم وما إلى ذلك. ومن قائمة جميع الخريجين تسحب عينة عشوائية في نقاط زمنية مختلفة ويتم جمع البيانات من تلك العينة. وهكذا يبقى المجتمع الإحصائي نفسه خلال الدراسة، ألا أن الأفراد الذين جرى المسح عليهم يختلفون في كل مرة.

المسوحات العرضية Cross-Sectional Surveys

تدرس المسوحات العرضية مقطعاً عرضاً (عينة) من مجتمع إحصائي في نقطة زمنية واحدة. ففي الدراسة الطولية لتطور المفردات، مثلاً، يقارن الباحث قياس مهارات مفردات طلبة الصف الأول عام 1990 مع طلبة من الصف الرابع عام 1993 وطلبة الصف السابع عام 1996. وتقارن الدراسة العرضية مهارات المفردات لعينة من الأطفال من الصف الأول والرابع والسابع عام 1996. وتعد الدراسة العرضية طريقة اختيار إن أراد أحدهم جمع البيانات في النقطة الزمنية نفسها.

ويكمن العيب الكبير للطريقة العرضية في أن فروقات الصدفة بين العينات قد تتحيز بشكل خطير في النتائج. فقد يسحب أحدهم بالصدفة عينة من طلبة الصف الأول الذين يكونون أكثر نضجاً من المتوسط وعينة من الصف الرابع الأقل نضجاً من المتوسط، مع نتيجة أن الفرق بين المجموعات يبدو أصغر مما هو عليه حقاً. وعلى أية حال، من المحتمل الحصول على عينات أكبر للدراسات العرضية مما يمكن الحصول عليه في الدراسات الطولية، فالعينات الكبيرة تقلل من مشكلة فروقات الصدفة.

أسلوب المسح THE SURVEY THECHNIQUE

يسمح المسح بجمع المعلومات من عينة كبيرة من الناس بشكل سريع ورخيص نسبياً. ومع ذلك، فإن إجراء المسح الجيد ليس بالأمر الهين كما يبدو أساساً. إنه يتطلب تخطيطاً دقيقاً، وتنفيذاً، وتحليلاً إن أريد له أن يوفر معلومات موثوقة وصادقة. وهناك. خمس خطوات أساسية في البحث المسحي:

1. **التخطيط** يبدأ البحث المسحي بسؤال يعتقد الباحث انه يمكن الإجابة عليه بشكل أكثر ما يكون مناسباً بواسطة طريقة المسح. إن أسئلة مثل كيف يشعر معلمو الابتدائية بشأن إبقاء الطلبة بعد الدوام؟ وما هو مدى استخدام التبغ بين طلبة الثانوية في هذه المقاطعة؟ هي أسئلة يمكن الإجابة عليها عن طريق المسح. إن السؤال البحثي في البحث المسحي يتعلق بشكل نموذجي بالمعتقدات أو التفضيلات أو الاتجاهات أو تقارير السلوك الذاتي التي يقدمها أفراد الدراسة عن أنفسهم. وستكشف مطالعة الأدبيات ما عرفه الباحثون الآخرون عن السؤال. ففي مرحلة التخطيط من المسح، يعتبر من المهم جداً تحديد المجتمع الإحصائي المستهدف والمجموعة التي يريد الباحث أن يعمم استنتاجات المسح عليها. (مثلاً، قد يكون المجتمع الإحصائي

المستهدف جميع معلمي الابتدائية في الولاية أو جميع طلبة الثانوية في مقاطعة مدرسية خاصة). ويحتاج الباحث أيضاً إلى القرار بشأن أسلوب جمع البيانات الذي سيتم استخدامه.

2. **المعاينة** تعتبر المعاينة، ذات اهتمام رئيسي في المسوحات. فينبغي على الباحث اتخاذ قرارات حول إجراء المعاينة الذي سيستخدم، وحجم العينة التي سيتم مسحها. وإذا تعين على أحدهم تعميم استنتاجات العينة على مجتمع إحصائي، فمن المهم أن تكون العينة المنتخبة ممثلة لذلك المجتمع. إن إجراء المعاينة الذي يمكن أن يتمخض عن عينة ممثلة، هو شكل معين من المعاينة الاحتمالية (انظر الفصل 6). وتسمح المعاينة الاحتمالية بتقدير مدى انحراف نتائج العينة عن قيم المجتمع الإحصائي.

3. **بناء الأداة** إن مهمة رئيسية في البحث المسحي، هي بناء الأداة التي ستستخدم لجمع البيانات من العينة. أما النوعان الأساسيان لأدوات جمع البيانات فهما المقابلات والاستبيانات.

4. **إجراء المسح** عندما يتم إعداد أداة جمع البيانات، ينبغي أن تخضع لاختبار ميداني لتحديد ما إذا كانت ستوفر البيانات المطلوبة. وتشمل هذه الخطوة أيضاً تدريب مستخدمي الأداة أو مقابلة الأفراد أو توزيع الاستبيانات عليهم والتحقق من دقة البيانات التي تم جمعها.

5. **معالجة البيانات** الخطوة الأخيرة تشمل ترميز البيانات، والتحليل الإحصائي، وتفسير ما ينتج عنها، ورفع تقرير بالنتائج.

إن العديد من الاعتبارات مشمولة بتنفيذ الخطوات السابقة. ولتحقيق التوازن فإن هذا الفصل يناقش هذه الاعتبارات بالتفصيل.

أساليب جمع البيانات

DATA-GATHERING TECHNIQUES

هناك طريقتان رئيسيتان تجمع فيهما البيانات في البحث المسحي: المقابلات، والاستبيانات. ولكل طريقة خياران مما يوفر أربع طرق مختلفة لجمع البيانات:

1- المقابلة الشخصية.

2- المقابلة الهاتفية.

3- الاستبيان البريدي.

4- الاستبيان المباشر.

وحيث أن جميعها تستخدم طريقة طرح الأسئلة فإن لكل منها خواص فريدة، ومزايا، ومساوئ يحتاج المرء إلى أخذها بنظر الاعتبار قبل اتخاذ القرار حول نوع الأداة المراد استخدامها.

المقابلات الشخصية Personal Interviews

في المقابلات الشخصية، يقرأ المقابل الأسئلة للمستجيب في وضع، وجها - لوجه، مع تسجيل الإجابات. إن أحد أهم الأوجه للمقابلة هو مرونتها. فللمقابل فرصة رصد الفرد والوضع الكلي الذي يستجيب فيه المرء. ويمكن تكرار الأسئلة أو إيضاح معانيها في حالة عدم فهم المستجيب لها، ويمكن للمقابل أن يضغط للحصول على معلومات إضافية عندما تبدو الاستجابة غير كاملة أو لا علاقة لها البتة.

يعتبر معدل الاستجابة الكبير مزية واضحة أخرى للمقابلة. ويشير "معدل الاستجابة" إلى نسبة العينة المنتخبة ممن يوافقون على إجراء المقابلة أو يعيدون الاستبيان كاملاً. وتكون معدلات الاستجابة في المقابلات، عالية جداً - ربما 90% أو أفضل.

ويزيد الاتصال الشخصي من احتمال قيام الفرد بالمساهمة وتوفير المعلومات المطلوبة. أما مع الاستبيانات البريدية، فإن الاتصال الشخصي يكون مفقوداً أو يحتمل أن يرفض الناس التعاون. ويؤدي ذلك إلى العديد من "عدم الإعادة" (أناس لم يكملوا الاستبيان ويعيدوه). إن معدل الاستجابة المتدني النموذجي للاستبيان البريدي (أقل من 30% على وجه العموم) لا يقلص حجم العينة فحسب، بل قد يتحيز في النتائج (Fowler, 1988). وعلاوة على ذلك، فإن بوسع المقابل الحصول على إجابة على كل أو معظم الأسئلة. وتمثل البيانات المفقودة مشكلة خطيرة للاستبيان البريدي.

ثمة فائدة أخرى هي الضبط الذي يتوفر للمقابل على السياق الذي تؤخذ فيه الأسئلة بنظر الاعتبار. وفي بعض الحالات يكون من المهم ألا يعرف المستجيبون طبيعة الأسئلة التالية لأن ردودهم على هذه الأسئلة قد تؤثر على الردود الأولى. ويتم إقصاء هذه المشكلة في المقابلة حيث لا يعرف الفرد الأسئلة التي ستطرح ولا يمكن له أن يعود ويغير الإجابات المعطاة سابقاً. أما بالنسبة للأشخاص الذين لا يستطيعون قراءة وفهم الاستبيان المكتوب، فإن المقابلات توفر الأسلوب الوحيد الممكن لجمع المعلومات

السيئة الرئيسية للمقابلة الشخصية تكمن في كونها أكثر كلفة من طرق المسح الأخرى. فاختيار المقابلين وتدريبهم وسفرهم إلى موقع المقابلة يجعل هذا الإجراء مكلفاً. فالأمر يستغرق وقتاً كثيراً للاتصال بالمستجيبين المحتملين وتحديد المواعيد وإجراء المقابلة فعلاً.

والسيئة الأخرى، هي احتمال تحيز المقابل حيث يحدث هذا عندما تؤثر مشاعر واتجاهات المقابل أو جنسه، أو عرقه وخصائصه الأخرى على الطريقة التي تطرح فيها الأسئلة أو تفسر. فالمقابل قد يشجع شفوياً أو غير شفوي أو يكافئ " الردود " الصحيحة التي تناسب توقعاته. ففي تحليل للردود المقدمة لبعض المسوحات الخاصة بالآراء حول الإجهاض، وجد أن جنس المقابل قد يكون أحد العوامل. فالنساء اللاتي تتكلمن مع نساء والرجال الذين يتحدثون مع

مقابلين ذكور يميلون إلى إعطاء ردود مؤيدة للاختيار على الأسئلة أكثر مما تفعل النساء اللاتي تتحدث مع رجال أو رجال يتحدثون مع نساء (Coughlin, 1990, Feb. 7).

ثمة مشكلة أخرى هي "تحيز الرغبة الاجتماعية" التي يريد فيها المستجيبون أن يبعثوا السرور لدى المقابل من خلال إعطائه ردوداً مقبولة اجتماعياً قد لا يعطوها بالضرورة في استبيان مجهول الاسم. مثلاً، في استفتاءات التفضيل في انتخابات عام 1989 التي تشمل مرشحي الأقليات لمنصب عمدة نيويورك ولمنصب حاكم فرجينيا، كانت نسبة المستجيبين ممن قالوا إنهم سيصوتون لصالح مرشحي الأقليات أعلى بكثير من نسبة الأصوات التي حصل عليها هؤلاء المرشحون في الانتخابات فعلاً. ولأخذ مثل هذا الخطأ بنظر الاعتبار، تصور الباحثون أنه ربما كان أصحاب الأصوات البيض يخشون أن الأمر قد يبدو عنصريا إن اعترفوا للمقابلين أنهم كانوا يفضلون مرشحاً أبيض (Couphlin, 1990, Feb. 7).

المقابلات الهاتفية Telephone Interviews

لقد أصبحت المقابلات الهاتفية أكثر شعبية حيث توضح الدراسات الحديثة على أنها تحظى باستحسان، مقارنة مع المقابلات المباشرة وجهاً-لوجه ((Wilhoit & Weaver, 1990. أما فوائدها الرئيسة فهي الكلفة المتدنية والإكمال السريع مع معدلات استجابات عالية نسبياً. ويمكن إجراء المقابلات الهاتفية على مدى وقت قصير نسبياً مع أشخاص منتشرين على مساحة جغرافية واسعة. مثلاً، تستخدم منظمات الاستفتاء القومي الهاتف في الغالب للحصول على آراء الناس في جميع أنحاء البلاد ممن سيصوتون قرب وقت الانتخابات. وغالباً ما تستخدم المسوحات الواسعة في المدن الكبيرة، الهاتف بدلاً من إرسال المقابلين إلى مناطق غير آمنة. فالهاتف يسمح بالوصول إلى أناس قد لا يفتحون أبوابهم لمقابل، غير أنهم قد يكونون راغبين في الحديث عبر الهاتف. وكفائدة أخرى هي أن للمستجيبين اعتقادا أكبر ببقائهم مجهولين – ومن هنا قد يكون هناك تحيز أقل لدى المقابل وتحيز أقل في الرغبة الاجتماعية مما هو عليه في المقابلات الشخصية.

إن للمقابلة الهاتفية أيضاً مزية أخرى، وهي إمكانية استخدام الحواسب في إجراء المقابلة وترميز الردود. فارتداد سماعات الأذن يجعل المقابل يجلس إلى الحاسوب بينما يختار هذا عشوائياً الرقم الهاتفي ويطلب الرقم. وعندما يرد المستجيب يقرأ المقابل الأسئلة التي تظهر على الشاشة ويطبع الأجوبة مباشرة في الحاسوب. وهذا يوفر الوقت للباحث الذي عادة ما يصرفه في الترميز والحصول على البيانات المنظمة، وإدخالها في الحاسوب لغرض التحليل.

أما السيئة الرئيسية للمقابلة الهاتفية، فهي وجود فرصة أقل لإقامة علاقة وئام مع المستجيب قياساً بالوضع الفعلي وجهاً – لوجه. ويتطلب الأمر مهارة كبيرة لتنفيذ المقابلة الهاتفية بحيث يمكن الحصول على نتائج صادقة. ويكون من الصعب عموماً، التغلب على الشكوك لدى المستجيبين المندهشين، خصوصاً عندما تطرح أسئلة حساسة أو شخصية. وقد تستخدم أحياناً

رسالة تُشعر المستجيبين المحتملين بالمقابلة القادمة وذلك لغرض التعامل مع المشكلة، إلا أن الرسالة قد تؤدي إلى مشكلة أخرى. فيكون لدى المتلقي عندئذ الوقت ليفكر بالردود أو ليعد رفضاً للمشاركة عندما تأتي المكالمة.

وثمة سيئة أخرى هي أن بعض البيوت ليس لديها هواتف والبيوت التي لديها أرقام غير مدرجة تستثنى تلقائياً من المسح، وهذا قد يؤدي إلى تحيز النتائج. على أية حال هناك أسلوب يعرف (بالاتصال الرقمي العشوائي) الذي يحل مشكلة الأرقام غير المسجلة (رغم أن ذلك لا يساعد في الوصول إلى البيوت التي ليس لديها هاتف). ففي الاتصال الرقمي العشوائي، يقدم الحاسوب عشوائياً قائمة بأرقام الهواتف تستند إلى جميع الأرقام المحتملة التي يعتقد أنها في الخدمة في منطقة ما. وبسبب التحديد العشوائي، فإن هذا الأسلوب يؤكد أن لكل بيت ذي خدمة هاتفية فرصة متساوية في شموله بالعينة. انظر (Dillman, 1978) لمناقشة عميقة حول الاتصال الرقمي العشوائي.

الاستبيانات البريدية Mailed Questionnaires

إن الاتصال المباشر مع الأشخاص في مقابلة شخصية يستهلك وقتاً كثيراً وتكاليف باهظة. ويمكن الحصول على المعلومات ذاتها، في الغالب، بواسطة استبيان يرسل بالبريد لكل فرد في العينة مع طلب إكماله وإعادته في وقت معين. ونظراً لأن الاستبيان يبعث بالبريد، فإن من الممكن شمول عدد كبير من الأفراد وكذلك لأفراد في مناطق متنوعة أكثر مما هو في المقابلة.

إن للاستبيان البريدي مزية ضمان السرية أو إغفال الاسم، مما يؤمن ردوداً أكثر صدقاً مما يمكن الحصول عليه في مقابلة شخصية. ففي المقابلة قد يحجم الأفراد عن التعبير عن آراء غير شعبية أو خاطئة سياسياً أو إعطاء معلومات يعتقدون أنها قد تستخدم ضدهم في وقت لاحق. ويقصي الاستبيان البريدي أيضاً مشكلة تحيز المقابلة.

أما سيئة الاستبيان فهي إمكانية سوء تفسير المستجيبين للأسئلة. فمن الصعب جداً صياغة سلسلة من الأسئلة تكون معانيها واضحة جداً لكل قارئ. فالباحث قد يعرف بالضبط ما يعنيه السؤال، غير أنه بسبب الصياغة الضعيفة أو المعاني المختلفة للمصطلحات، قد يجد تفسيراً بالغ الاختلاف لدى المستجيب. وعلاوة على ذلك، فإن شرائح كبيرة من السكان قد لا تكون قادرة على القراءة والرد على استبيان بريدي. فالأفراد ذو الثقافة العالية فقط، قد يكونون قادرين على إكمال استبيان معقد جداً.

وثمة قصور مهم آخر للاستبيان البريدي هو معدل الإرجاع المتدني. فمن السهل بالنسبة للفرد الذي يتلقى الاستبيان أن يضعه جانباً وينسى إكماله وإعادته. فمعدل الرد المتدني يحد من إمكانية تعميم نتائج دراسة الاستبيان. فلا يمكن الافتراض بأن غير المستجيبين يتوزعون عشوائياً في المجموعة كلها. فقد أوضحت الدراسات بأن هناك فروقات منتظمة عادة في خصائص

المستجيبين وغير المستجيبين في دراسات الاستبيان. إن معدل الاستجابة هو في الغالب أعلى بين الأكثر ذكاء، والأفضل ثقافة، والأكثر وعياً، والأكثر اهتماماً، والأكثر تفضيلاً للقضية المعنية في الاستبيان. والهدف في دراسة الاستبيان هو إعادة الردود (100%)، رغم أن التوقع المعقول هو (75-90%).

لقد تبين أن عدداً من العوامل تؤثر على معدل إعادة الاستبيان البريدي. وبعضها هو

(1) طول الاستبيان، (2) رسالة الغلاف / التفسير، (3) دعم الاستبيان، (4) جاذبية الاستبيان، (5) سهولة إكماله وإرجاعه بالبريد، (6) الاهتمام الذي يثيره المحتوى، (7) استخدام الحافز المالي، (8) إجراءات المتابعة المستخدمة.

وستناقش هذه بتفصيل أكبر في الأجزاء التالية.

الاستبيانات المباشرة

Directly-Administered Questionnaires

إن الاستبيان المباشر هو استبيان يجري تسليمه لمجموعة من الأفراد يجتمعون في مكان معين ولهدف محدد. وتشمل الأمثلة مسح الطلبة الجدد، أو أولياء أمورهم الذين ينتظمون في دورة توجيه / إرشاد صيفية في جامعة معينة. وتجري المسوحات في الجامعات، غالباً، في الصفوف أو في السكن الداخلي. لقد استخدم (Jacobs, 1985). هذه الطريقة لجمع إدراك الطلبة الجدد لما يتعلق بالإعداد الأكاديمي لعمل الكلية. وكان من السهل الوصول إلى عينة كبيرة من الطلبة في حقول مختلفة عن طريق إجراء مسح الصفوف (بإذن من الأساتذة).

وتكمن المزية الرئيسية للتسليم المباشر للاستبيان في معدل الاستجابة العالي الذي يصل (100%). هناك مزايا أخرى مثل الكلفة المتدنية وحقيقة أن الباحث موجود لتقديم المساعدة أو الإجابة على الأسئلة. أما السيئة فهي أن الباحث مقيد عادة في إطار متى وأين يتم تسليم الاستبيان. وكذلك بسبب كون أن العينة محددة في العادة (مثلاً، أولياء أمور طلبة جدد في جامعة محددة) فإن الاستنتاجات قابلة للتعميم فقط على المجتمع الذي تمثله العينة.

SELECTING THE SAMPLE اختيار العينة

نشجع القارئ على العودة مرة أخرى إلى الجزء في الفصل (6) الخاص بأساليب المعاينة والعوامل التي تؤثر على حجم العينة المنتخبة. ويتبين بعض النقاش الإضافي حول حجم العينة في الجزء التالي.

حجم العينة Sample Size

إن مدى كبر حجم العينة هو أحد الأسئلة الأولى التي يجب على الباحث الإجابة عليها. فكيف يقرر المرء بشأن عدد الناس الذين يراد مسحهم؟ ويعتقد العديد من الباحثين أن عليهم اختيار عينة تكون على الأقل (10%) من المجتمع الإحصائي، لكن ذلك غير ضروري. وعلى عكس ما هو سائد عموماً، فإن دقة البيانات تتحدد "بالحجم المطلق" للعينة وليس بالنسبة المئوية لحجم العينة من المجتمع. وقد تلاحظ من استفتاءات الرأي العام الرئيسة في الولايات المتحدة أنها لا تستخدم أعدادا كبيرة. مثلاً، مجلة نيوزويك، في مسح قومي عام 1994 حول الهوموفوبيا (الخوف من الأجانب) أجرت مقابلات مع (750) فردا - وهو عدد صغير بالنسبة لحجم المجتمع الإحصائي.

إن الاعتبار الرئيسي عند اتخاذ القرار حول حجم العينة هو درجة الدقة التي يريدها المرء في تقدير قيم المجتمع. فما مقدار الخطأ الذي يشاء الباحث أن يتحمله في تعميمه من إحصاءة العينة إلى معْلمة المجتمع الإحصائي(*)؟ فإذا استخدم الباحثون المعاينة الاحتمالية، يكون لديهم أساس في تقدير احتمال مدى انحراف نتائج العينة عن قيم المجتمع - أي هامش الخطأ- بالنسبة لحجم معين لعينة. وهم يختارون حجم عينة تمكنهم من أن يكونوا واثقين بان تقديراتهم ستكون صحيحة ضمن مدى صغير في حوالي (95%) من الوقت. دعنا نوضح ذلك.

تقدير قيم المجتمع الإحصائي
Estimating The Population Values

في حالة المتغيرات ذات الحدين (قيمتان فقط) فإن التوزيع ذو الحدين يقدم أساساً لتقدير نسبة المجتمع التي يتمتع أفرادها بخاصة معينة. مثلاً، ما نسبة أصحاب الأصوات المسجلين الذين يفضلون تقليص ميزانية الدفاع القومي؟ أو ما نسبة أصحاب الأصوات الأمريكيين الذين يدعمون تغطية الرعاية الصحية؟ إنا نستخدم الردود من العينات لتقدير نسب المجتمع. فعندما يكون (N) كبيرا (50 أو أكثر) فإن نسب العينات تتوزع بشكل اعتيادي تقريباً، مع وسط حسابي يساوي "P" (نسبة المجتمع) وانحراف معياري يدعى بالخطأ المعياري له المعادلة التالية:

$$\text{الخطأ المعياري} = \sqrt{\frac{Pq}{N}}$$

حيث:

p = النسبة التي تقدم استجابة واحدة (مثل، نعم).

(*) نود التذكير أن قيمة إحصائية للعينة تسمى "إحصاءة Statistic" والمقابلة لها في المجتمع الإحصائي تسمى "معْلمة / معْلم : Parameter"، وقد أشرنا إلى ذلك في هوامش متعددة للتذكير المفيد – (المراجع).

q = النسبة التي تقدم الاستجابة الأخرى (كلا).

pq = تباين العينة.

N = حجم العينة.

إنك تذكر أن الخطأ المعياري هو مقياس دقة بيانات العينة كتقدير لقيمة المجتمع. فكلما كان الخطأ المعياري صغيراً كان هناك احتمال أكبر بتمثيل العينة للمجتمع. افترض أننا قمنا بمسح عينة عشوائية من (500) ناخب وسئلوا ما إذا كانوا يفضلون تقليص ميزانية الدفاع فأجاب (375) منهم أو (75%) بالقول (نعم) و (125) أو (25%) قالوا (كلا). هل يمكننا أن نعتبر أن (.75) تقدير للنسبة الحقيقية لجميع الناخبين ممن يفضلون تقليص ميزانية الدفاع؟ إن مدى جودة التقدير يمكن تحديده عن طريق وضع حدود الثقة للمجتمع الحقيقي. فكلما كانت هذه الحدود قريبة من (.75) كان التقدير أدق. وتكون الخطوة الأدنى هي حساب الخطأ المعياري:

$$s.\ e. = \sqrt{\frac{(.75)(.25)}{500}} = .0194$$

نستخدم الخطأ المعياري لإقامة فترة يحتمل لها أن تحتوي على القيمة الحقيقية للمجتمع كنسبة مئوية لذلك الوقت. وعند مستوى الثقة البالغ (95%) الذي يعتبر المستوى التقليدي، تكون الفترة هي قيمة العينة (+1.96) خطأً معيارياً.

تذكر من الفصل (5) أن (95%) من المنطقة تحت المنحني الاعتيادي يقع بين (+1.96) انحرافاً معيارياً و (−1.96) انحرافاً معيارياً عن الوسط الحسابي. وفي هذه الحالة تكون فترة النسبة المئوية للأفراد في المجتمع الذين يفضلون تقليص ميزانية الدفاع:

$$.75 \mp 1.96\ (.019) = .75 \mp .037$$

أي أن هناك احتمالاً مقداره (.95) بأن النسبة الحقيقية في المجتمع ممن يفضلون ميزانية دفاع متدنية هو ما بين 71% −79% (.75 ∓ .037). وهكذا فإن هامش الخطأ لهذا المسح هو (4) نقاط مئوية. وإن أراد الباحث هامش خطأ صغير تكون هناك حاجة لعينة أكبر. وبوسع المرء أن يرى من المعادلة الخاصة بالخطأ المعياري أن زيادة (N) تقلص من حجم الخطأ المعياري. فإذا جرى في الدراسة أعلاه مسح (1500) ناخب بدلاً من (500) فإن الخطأ المعياري سيتقلص إلى:

$$\sqrt{\frac{(.75)(.25)}{1500}} = .011$$

إن حدود الثقة البالغة (95%) ستمتد فقط إلى [(1.96×.011) ∓ .75] ∓.02 معطياً (73) إلى (77) كفترة ثقة يمكن لها أن تحتوي قيمة المجتمع. فزيادة حجم (N) زادت دقة التقدير لهامش الخطأ من (4) إلى (2) نقطة مئوية. فما مقدار الحجم الذي يجب أن تزداد به (N) بغية

تحقيق مستوى معين من الدقة؟ افترض أن مدير مدرسة أراد أن يعرف كيف يشعر الطلبة حول حصولهم على برنامج تربوي إجباري حول مرض الإيدز في المدرسة. ويود المدير أن يكون التقدير دقيقاً إلى حد (4) نقاط مئوية عند مستوى الثقة (95%). وبالوسع استخدام المعادلة لتحديد فترة الثقة (1.96 × s.e.=.04) لإيجاد (N). والمشكلة الوحيدة هي أن المدير لا يعرف (P) في معادلة الخطأ المعياري. وبوسع المدير أن يقوم بأحد شيئين: أولا يمكن أن يجري دراسة استطلاعية مع (N=25) بغية الحصول على تقدير (P). افترض أنه وجد أن (7) من (25) طالبا يقولون أنهم يفضلون البرنامج (P =7/25 =.28). فإذا كانت (P= .28) فعندئذ (q = .72). وباستبدال ذلك بالمعادلة:

$$1.96\sqrt{\frac{pq}{N}} = .04$$

فيكون لدينا

$$1.96\sqrt{\frac{(.28)(.72)}{N}} = .04$$

ويمكن إيجاد (N) بتقسيم كلا الطرفين على (1.96) وتربيع كليهما:

$$\sqrt{\frac{(.28)(.72)}{N}} = \frac{.04}{1.96}$$

$$\frac{(.28)(.72)}{N} = \left(\frac{.04}{1.96}\right)^2$$

$$\frac{.2016}{N} = .0004164$$

$$N = 484$$

قد يحتاج المدير إلى عينة من (484) طالبا بغية إقامة فترة ثقة بمقدار (95%) وهي (.04 ∓ .28). افترض أن المدير يجري مسحاً على (484) طالبا ويجد أن (121) طالبا يقولون الهم يريدون البرنامج، فتكون نسبة العينة الآن هي (121/484) .25. فبوسعه أن يستخدم (P = .25) ويحل المعادلة لإيجاد (N) مرة أخرى:

$$1.96\sqrt{\frac{(.25)(.75)}{N}} = .04$$

هذه المرة (N=450) مما يقنع المدير أن العينة الأصلية من (484) طالبا كانت أكثر من كافية للتقدير (P=.25). وبوسع المدير أن يقول أن (25%) من الطلبة في المدرسة يفضلون

البرنامج التربوي حول مرض الإيدز. أما هامش الخطأ فهو (4 ∓) نقاط مئوية. أما الاحتمال فهو (95%) بأن تحتوي الفترة (04.∓ 25.) على قيمة حقيقية للمجتمع الإحصائي. وهكذا يكون المدير قادراً أيضاً على حساب حجم العينة المستخدمة للحصول على حدود الثقة والخطأ المرغوبين. وإذا أدى الحساب الثاني لـ (N) إلى قيمة أكبر من (484) فإن من الضروري زيادة عدد الطلبة الذين جرى مسحهم ثم إعادة الحساب.

ثمة إجراء آخر يمكن للمدير استخدامه مستبعداً الدراسة الاستطلاعية ويكون بمجرد الافتراض أن (P= .50). وبحل المعادلة:

$$1.96\sqrt{\frac{(.50)(.50)}{N}} = .04$$

فتكون : (تقريباً) N = 600

إن استخدام (P=.50) هو طريقة آمنة دائماً لحساب حجم العينة لأن (pq) يكون في أقصى قيمة محتملة لها عندما (p=q=.50)، والنتيجة هي أكبر تقدير ممكن لحجم العينة المطلوب. ومن ثم، إذا وجد المرء أن (P) هو أكبر أو أصغر من (.50) فإن الحصول على الحجم الأكبر للعينة المحسوب بافتراض أن (P=.50) سيعطي تقديراً أدق مما يتوقع. وهذا الإجراء هو الطريقة الأكثر تحفظاً، إن لم تكن الأكثر اقتصاداً للوصول إلى حجم العينة. ويبين الجدول 12.1حجم العينة المطلوب للحصول على هامش خطأ معيّن لـ (P). لاحظ إن العينة الأكبر تكون مطلوبة عندما يكون هامش الخطأ المسموح به هو الأدنى وأن (P= .50). وعندما يزداد هامش الخطأ المقبول وتختلف (P) عن (.50) يتقلص حجم العينة المطلوب.

الجدول 12.1: الحجوم الصغرى للعينات المطلوبة لهوامش أخطاء حول تقدير المعْلمة (parameter) حسب مستوى الثقة (.95).
(الحجوم موضحة كدالة لنسب العينات المتوقعة)

قيمة P			أقصى هامش خطأ
.50	.75 أو.25	*.90 أو .10	
9616	7212	3462	1%
2404	1803	866	2%
1069	802	385	3%
385	289	139	5%
97	73	39	10%

* يظل تباين العينة (pq) على حالة عندما (.90) (.10) = pq أو عندما (.10) (.90) = pq

إن زيادة حجم العينة يمكن أن يكون باهظ التكاليف. فإرسال الاستبيانات بريدياً والمتابعات أمر مكلف، كما أن إجراء المقابلة أكثر كلفة. فيجب على الباحث أن يأخذ بالاعتبار الوقت والمال المتيسر للمسح وينتخب أكبر عينة ممكنة بالقدر الذي تسمح به الحالة الاقتصادية. لكن عليك أن تتذكر: الحجم وحده لا يضمن عينة ممثلة، فإجراء المعاينة أكثر أهمية في تحديد ما إذا كانت العينة ممثلة للمجتمع الإحصائي. ونشجع القارئ على قراءة الجزء في الفصل (6) حول إجراء المعاينة والعوامل الأخرى لأخذها بالاعتبار عند اتخاذ القرار حول حجم العينة.

بنـــاء الأداة
CONSTRUCTING THE INSTRUMENT

أنواع الأسئلة Types of Questions

نظراً لأن بيانات المسح تتكون من أجوبة الناس على الأسئلة، فمن المهم جداً البدء بأسئلة جيدة. ويستخدم نوعان رئيسان من الأسئلة في أدوات المسح: "مغلق النهاية" أو "مفتوح النهاية". فيستخدم أحدهم الأسئلة مغلقة النهاية عندما يمكن تحديد جميع الردود المحتملة ذات الصلة بسؤال معين ويكون عدد الردود المحتملة محدوداً. مثلاً، في مسح لطلبة المرحلة الجامعية الأولى، يكون السؤال عن مستوى الصف مغلق النهاية. وتكون الأجوبة المحتملة معروفة وقليلة العدد: جديد، سنة ثانية، سنة ثالثة، سنة رابعة.. وأسئلة أخرى مغلقة النهاية قد تسأل عن وضع الإقامة (في الولاية أو خارجها)، أو الجنس. أما الأسئلة مفتوحة النهاية فتستخدم عندما يكون هناك عدد كبير من الأجوبة المحتملة، أو عندما لا يكون الباحث قادراً على التنبؤ بجميع الأجوبة المحتملة. مثلاً، سؤال عن أسباب الطلبة في اختيار جامعة معينة ربما يكون مفتوح النهاية. وسؤال حول التخصص الرئيسي في الكلية يكون مفتوح النهاية لأن الباحث قد لا يرغب أن يدرج قائمة طويلة للتخصصات الممكنة. ويمكن استخدام كلا الصيغتين في السؤال نفسه – أي يمكن إلحاق عدد من الاستجابات مغلقة النهاية بعبارة "آخر" كإجابة محتملة أخيرة. فمثلاً، قد تسأل أستاذا سؤالاً مثل:

1- ما نوع واجبات / تعيينات الكتابة التي تطلبها عادة في مقررك الدراسي؟ (ضع دوائر بالقدر الذي تراه مناسباً).

أ. تقارير

ب. موضوع أو مقالة

ج. أوراق بحثية

د. اختبارات مقالة بيتية

هـ. أوراق بحثية مصغرة
و. آخر (حدد رجاء) ____________________

ثمة مزايا ومساوئ لتصميم كل من السؤالين هذين. فالسؤال مفتوح النهاية يسمح باستجابة حرة ولا يقيد المستجيب باختيار من بين البدائل الموضحة. فالأفراد أحرار في الاستجابة ضمن إطار تفضيلاتهم مما يقدم مدى واسعا من الاستجابات. والأسئلة مفتوحة النهاية أسهل بناءاً، إلا أن تحليلها ممل ويستهلك وقتاً طويلاً. ويجب على الباحث أن يقرأ ويفسر كل استجابة، وأن يطور استخدام ترميز يجعل من الممكن إجراء تحليل كمي للاستجابات. وقد تكون بعض الاستجابات غير واضحة، وقد يتعذر على الباحث تصنيف أو ترميز الاستجابة. أما الاستجابات للأسئلة مفتوحة النهاية فتختلف، كما هو معهود، في طولها، وقد يعطي بعض المستجيبين أكثر من استجابة لسؤال معين. فالمستجيب الذي يُسأل عن أهم سبب لاختيار جامعة معينة قد يجيب: "اخترت الجامعة X بسبب سمعتها الأكاديمية ولأنها ضمن الولاية واقل كلفة بالنسبة لي". وعلى الباحث أن يقرر ما إذا كان سيستخدم كلا الإجابتين أو الإجابة الأولى على افتراض أن الإجابة الأولى هي الأهم.

أما الأسئلة مغلقة النهاية فإنها تأخذ وقتاً أكثر لبنائها إلا أن الاستجابات أكثر سهولة في وضعها وتنظيمها. ويمكن ترميز الاستجابات على الأسئلة مغلقة النهاية مباشرة على رقائق قابلة للتدقيق يمكن "قراءتها" ثم توضع البيانات في الحاسوب لغرض التحليل. ويمكن الإجابة على الأسئلة مغلقة النهاية بشكل أسهل وأسرع من قبل المستجيبين. كما أن البنية المغلقة تضمن أن يكون لجميع الأفراد إطار المرجعية ذاته في الاستجابة وقد تجعل من الأسهل على الأفراد، الإجابة على أسئلة تعالج موضوعات ذات طبيعة حساسة أو خاصة.

ويكمن موطن الضعف في السؤال مغلق النهاية في كونه لا يوفر الكثير من التبصّر فيما إذا كان لدى المستجيبين حقاً أية معلومات أو أية آراء مصاغة بشكل واضح حول قضية معينة. ومن السهل للمستجيب غير المطلع أن يختار أحد الأسئلة المقترحة بدلاً من الاعتراف بالافتقار إلى المعرفة حول قضية معينة. مثلاً، في الرد على السؤال "ما أهم الفوائد لاتفاقية التجارة الحرة في أمريكا الشمالية بالنسبة إلى الولايات المتحدة؟ "يمكن للمستجيب الذي ليس لديه معرفة كبيرة بالاتفاقية أن يختار بسهولة الجواب المعقول من بين البدائل المقدمة. من ناحية أخرى، قد ينفر المستجيبون، ممن لديهم معرفة أو آراء مطلقة حول قضية معينة، من التقيد بأنواع استجابات بسيطة لا تسمح لهم بإبراز جدارة إجاباتهم.

من الممكن الحصول على فوائد الأسئلة مفتوحة ومغلقة النهاية. وبوسع الباحث استخدام الصيغة مفتوحة النهاية مع عينة صغيرة بهدف تحديد الاستجابات البديلة المحتملة على الأسئلة. وعلى أساس النتائج من العينة، يمكن للباحث أن يصمم أسئلة مغلقة النهاية للشكل النهائي للأداة.

بنية الأسئلة Structure of Questions

ناقشنا مزايا ومساوئ الأسئلة مفتوحة ومغلقة النهاية في البحث المسحي. دعنا الآن ننظر إلى الطرق التي يمكن فيها استخدام الصيغتين لبناء أسئلة من أجل المقابلات والاستبيانات.

1- "إكمال وملء البنود" وهي أسئلة مفتوحة النهاية يجب أن يقوم المستجيبون بتقديم أجوبتهم بكلماقم. مثلاً:

ما هو الضعف الرئيسي الذي لاحظته في إعداد طلابك للكلية؟ ______________

____

2- "قوائم التدقيق" وهي أسئلة تعرض عدداً من الأجوبة المحتملة ويطلب من المستجيبين أن يدققوا تلك التي تنطبق. مثلاً:

ما نوع وسائل التعليم المعينة التي تستخدمها في صفوفك؟ (دقق أكبر عدد من الحالات التي تنطبق).

______ 1) سبورة.
______ 2) جهاز العرض العلوي.
______ 3) جهاز عرض حاسوبي.
______ 4) أشرطة فيديو.
______ 5) أخرى (حدد رجاءً). ______________________

3- "بنود مقيسة" تتطلب من المستجيبين وضع تقدير لمفهوم، أو حادث، أو موقف، وفق أبعاد، فحسب الكم أو الشدة، توضح المقدار"، وحسب النوع، توضح "المدى/ الدرجة"، وحسب التكرار توضح "عدد المرات" مثلاً:

كيف تصنف / تقدر مهارات الكتابة لدى الطلبة الذين تدرسهم هذا الفصل الدراسي؟ (دقق واحدة)

______ 1) ضعيف جداً.
______ 2) أقل من كاف.
______ 3) كاف.
______ 4) أكثر من كاف.
______ 5) ممتاز.
______ 6) معلومات غير كافية.

أو ما مدى التهيؤ في مهارات الرياضيات الأساسية لدى الطلبة الذين يسجلون، كما هو معهود، في مقررك؟ (دقق واحدة).

______ 1) غير مهيئين البتة.
______ 2) مهيئون نوعا ما.
______ 3) مهيئون بشكل جيد جدا.
______ 4) مهيئون على نحو بالغ.

4- "أسئلة الترتيب" تسأل المستجيبين إيضاح ترتيب تفضيلهم بين عدد من الخيارات. ولا ينبغي للترتيب أن ينطوي على أكثر من ستة خيارات، لأنه يصبح من الصعب على المستجيبين إجراء المقارنات. ومثال عن بند للترتيب هو كما يلي:

هل يواجه طلبتك صعوبة أكثر في بعض أنماط القراءة مما عليه الحال في أنماط أخرى؟ يرجى ترتيب الصعوبة في المواد التالية، مع (1) للأكثر صعوبة و (4) للأقل صعوبة.
______ كتب مقررة.
______ كتب مرجعية أخرى.
______ مقالات مجلات.
______ أخرى (حدد رجاءً).

5- "بنود من نمط ليكرت" وهذه تتيح للأفراد توضيح استجاباتهم إزاء فقرات مختارة في متصل كمي (continuum) من موافق بشدة إلى معارض بشدة. وتكمن مزية هذا النوع من البنود هو إمكانة تعيين نقاط / درجات للاستجابات المختلفة، لذا يمكن عندئذ حساب مقاييس النـــزعة المركزية، والتباين، والارتباط وما شابه. مثلاً:

الطلبة الذين يسجلون عادة في مقرري هم دون التهيؤ في مهارات الرياضيات الأساسية.

أوافق بشدة   أوافق   غير مقرر   لا أوافق   لا أوافق بشدة

كتابة أسئلة المسح Writing Survey Questions

دعنا الآن نفحص مهمة صياغة أسئلة جيدة. فالقرار حول صياغة الأسئلة التي تطرح في مسح معين هو قرار صعب. فقبل البدء بكتابة مجموعة منظمة من الأسئلة المسحية فإنه من المفيد توفير مجموعات تركيز (focus groups) تناقش الأسئلة بصيغة غير منظمة (Fowler, 1988). ويقوم المرء بذلك بجمع مجموعات صغيرة من خمس إلى عشرة أشخاص يمثلون مجتمع الدراسة بغية مناقشة الموضوعات التي يشملها المسح. ويعمل وسيط (رئيس لحلقة النقاش) على حفظ النقاش مركزاً على جدول الأعمال المعد مسبقا ويوجه أسئلة لإيضاح التعليقات.

تساعد نقاشات مجموعات التركيز الباحث على فهم الكيفية التي يتحدث بها الناس عن قضايا المسح، مما يساعد في اختيار المفردات وصياغة الأسئلة. وغالباً ما يكون بوسع مجموعات التركيز اقتراح قضايا، واهتمامات، ووجهات نظر حول موضوع معين لم يأخذه الباحث بنظر

الاعتبار. لقد أستخدم (Emery et al., 1993) مجموعات التركيز بغية تحديد قضايا بارزة تتعلق بإساءة تناول طلبة الكلية للكحول. وأقام الباحثون أربعة مجموعات تركيز لطلبة تم اختيارهم عشوائيا – واحدة كلها رجال، واثنتان تضمان نساءً فقط، وواحدة فيها رجال ونساء. ناقشت المجموعات أسباب تناول الطلبة للكحول، وأسباب عدم تناولهم للكحول، وطرق منع تناولها، وفروقات الجنس في أنماط الشرب. لقد أستنتج الباحثون أن مجموعات التركيز كانت أداة صادقة لتوضيح الجوانب الحساسة للقضايا والأهمية النسبية لهذه القضايا إزاء بعضها بعضا. وتضم المصادر الممتازة عن المعلومات المعمقة حول استخدام مجموعات التركيز (Morgan, 1988) وأيضاً (Stewart & Schamdasani, 1990).

تبين بحوث كثيرة أن التغيرات في أمور مثل الصياغة، ومقدار المعلومات المقدمة، واختيار الأجوبة المتوفرة للمستجيبين يمكن لها أن تؤثر على نتيجة المسح بدرجة كبيرة أو قليلة. ففي مقابلة مع (Coughlin, 1990, Feb. 7)، قال Ballou إن "إحدى القواعد في هذا العمل هو أن السؤال يحدد الجواب" (ص A6). وفيما يلي بعض الإرشادات الأساسية:

1- **ينبغي أن تكون الأسئلة قصيرة وبسيطة ومباشرة.** احذف أية كلمات أو عبارات غير جوهرية للمعنى الواضح للسؤال. فالأسئلة القصيرة أيسر فهما. والقاعدة المفيدة الناشئة عن التجربة ترى ضرورة أن تضم أغلب الأسئلة أقل من عشر كلمات (سطر واحد)، كما ينبغي أن تكون جميع الأسئلة دون العشرين كلمة (Mitchel & Jolley, 1988).

2- **قم بصياغة الأسئلة بحيث يتسنى لكل مستجيب فهمها.** ينبغي أن تكون المفردات المستخدمة غير فنية / اصطلاحية ويجب أن توجه إلى المستجيب الأقل ثقافة. وفي الوقت ذاته، ينبغي أن يتجنب الباحث التعالي على المستجيبين أو اختيار كلمات تبدي مناصرة ما. إنها فكرة جيدة أن يكون هناك أناس آخرون ممن تكون خلفياتهم مشابهة للمشمولين بالدراسة، فيقرءون ويفسرون محتوى كل سؤال. مثلاً، الأسئلة التي تستخدم مصطلحات مثل "التقدير الصادق" و "التعلم عن بعد" و "إدارة الجودة الكاملة" قد لا تكون مناسبة في مسح مصمم للناس بشكل عام. كما ينبغي أن يكون الباحث دقيقاً في عدم استخدام اللهجة المحلية، والاختصارات أو الكلمات المنحوتة من أوائل الحروف التي قد لا تكون مألوفة لدى الجميع.

3- **قم بصياغة الأسئلة بحيث تحدد أجوبة لا لبس فيها.** فسؤال مثل "هل قمت بالتصويت في الانتخابات الأخيرة؟ "هو سؤال غامض لأنه لا يحدد أية انتخابات. ويجب أن يكون التعبير عن الاستجابات بصورة كمية متى أمكن ذلك. فلكلمات مثل "غالباً" و "أحيانا" معان مختلفة بالنسبة للناس المختلفين. مثلاً، في مسح عن عدد المرات التي يستخدم فيها طلبة الكلية المكتبة الرئيسية للدراسة، ينبغي أن تكون الاستجابات محددة بالكم (يومياً، خمسة أيام في الأسبوع، مرتان في الأسبوع وما إلى ذلك) بدلاً من استخدام استجابات مثل عادة و أحيانا ومراراً.

4- **قم بصياغة أسئلة بحيث تتجنب التحيز الذي قد يحدد مسبقاً جواب المستجيب.** ينبغي أن لا تؤثر صياغة السؤال على المستجيب في اتجاه معين. ولهذا السبب فإنه ينبغي تجنب الكلمات التي تصنف الناس حسب آراء مقولبة (stereotyped)، أو التي تحمل مضمون المكانة، أو المشحونة عاطفياً أو التي تستخدم التفضيل / المبالغة. وثمة بعض الكلمات تكون لها جاذبية عاطفية في الثقافة بما يدعو إلى تحيز الأسئلة بغض النظر عن الكيفية التي تستخدم فيها. مثلاً، السؤال "هل تمارس حقك الأمريكي وتسجل من أجل التصويت؟" قد يعمل دون شك على تحيز السؤال. لكن السؤال البسيط "هل سجلت للتصويت؟" سيكون مفضلاً. يقول (Dillman, 1978) أن كلمات مثل "حرية، مساواة، مشروع خاص، عدالة، نزاهة" لها جاذبية إيجابية قوية في ثقافتنا. وكلمات مثل "بيروقراطي، اشتراكي، مدير و تخطيط حكومي" لها جاذبية سلبية شديدة. فمثل هذه الكلمات يجب تجنبها قدر الإمكان. لقد أعطى (Smith, 1987) مثالاً جيداً عن مدى تأثير بعض الكلمات في السؤال على الاستجابة. ففي مسح اتجاهات الناس حول الإنفاق على الرعاية الاجتماعية، طرح سميث السؤال ذاته بطرق مختلفة. لقد سئل بعض الناس ما إذا كان ينبغي علينا أنفاق مال أكثر على "الرعاية الاجتماعية" بينما سئل آخرون ما إذا كان يتعين علينا أنفاق مال أكثر على "دعم الفقراء" فالسؤال الذي يستخدم "الرعاية الاجتماعية" أدى إلى ردود سلبية أكثر. فلكلمة "الرعاية الاجتماعية" معنى البيروقراطية الحكومية، والهدر، وأحيانا الاحتيال بالنسبة للعديد من الناس. وعبارة "دعم الفقراء" تنطوي على العناية، والعطاء، والإحسان.

5- **تجنب الأسئلة التي قد تكون مضللة بسبب فروض غير مذكورة.** ينبغي أن يكون الإطار المرجعي لإجابات الأسئلة واضحاً ومتّسقاً لكل المستجيبين. وإذا كانت هناك افتراضات لابد من توضيحها للمستجيبين، فإن الأسئلة المصممة لاستقصاء هذه الافتراضات يجب أن تكون مشمولة. مثلاً، في مسح مصمم لطلبة السنة النهائية في الثانوية، فإن السؤال "هل تعتقد أن المدرسة الثانوية قد أعدتك بشكل كاف للكلية؟ "يفترض أن الطالب يذهب إلى الكلية ولديه معرفة بما هو مطلوب في طريق الأعداد. أما السؤال "هل سجلت اسمك لتصوت في انتخابات الرئاسة التالية؟" فيفترض أن طالب الثانوية هو في الثامنة عشرة من العمر، وقد لا يكون ذلك صحيحا.

6- **تجنب الأسئلة الإيحائية.** مثلاً، " لماذا تفضل الإصلاح الشامل لنظام الرعاية الصحية في الولايات المتحدة؟ هو سؤال إيحائي.

7- **تجنب الأسئلة التي قد تثير ردود أفعال مثل الإحراج، أو الشك، أو العداوة لدى المستجيب.** لا ينبغي للأسئلة أن تضع المستجيب في موقف الدفاع. مثلاً، يستاء الناس في الغالب من الأسئلة عن العمر، أو الدخل، أو الدين، أو الوضع الثقافي، فبدلا من السؤال عن عمر شخص ما، يمكن للباحث أن يسأل عن سنة ميلاده. ويبدو الناس أقل تحفظاً بإعطاء

سنة ميلادهم بدلاً من إعطاء أعمارهم. والسؤال "هل لديك شهادة الدبلوم من الثانوية؟ قد تسبب الإحراج لشخص لم يتخرج من الثانوية. وقد يكون السؤال "ما الصف الذي أكملته عندما تركت المدرسة؟ وفي الواقع، من الأفضل تجنب الأسئلة الشخصية كلية ما لم تكن المعلومات جوهرية للباحث.

8- **تجنب الأسئلة المزدوجة (double-barreled) التي تسعى لطرح سؤالين في سؤال واحد.** فالسؤال "هل تشعر أن على الجامعة أن توفر مقررات في المهارات الأساسية للطلبة وتمنح اعتماداً لهذه المقررات؟ هو سؤال مزدوج. فعندما يجيب المستجيب على سؤال مزدوج، فإن الباحث لا يعرف ما إذا كان الجواب ينطبق على جزئي السؤال أو على جزء واحد فحسب. فالجواب (نعم) على السؤال أعلاه قد يعني أن المستجيب يعتقد أن على الجامعة أن تقدم مقررات في المهارات الأساسية وتعتمدها أو أنها يجب أن تعطي مقررات دون أن تمنحها الاعتماد. ويمكن تحديد السؤال المزدوج من خلال (و) أو (أو) أو أية كلمة اقتران أخرى.

9- **تأكد أن البدائل لكل بند في الاستبيان مستوفاة / شاملة – أي أنها تعبر عن جميع البدائل الممكنة في القضية.** مثلاً، "ما هي حالتك الزوجية؟" ينبغي ألا يشمل بدائل مثل متزوج وعازب فحسب، بل كذلك أرمل، مطلق، منفصل. وبتطوير بدائل بنود الاستبيان المصممة لتحديد الاتجاهات، أو الآراء حول قضية معينة، فإنها فكرة جيدة أن تقدم الأسئلة أولاً بطريقة مفتوحة إلى عينة صغيرة من المستجيبين. ويمكن استخدام أجوبتهم كبدائل في الحاصل النهائي. وحول الأسئلة ذات التنوع الواسع من الاستجابات المحتملة ينبغي أن يشمل المرء دائما البديل "آخر" إضافة إلى الطلب من المستجيب إيضاح ذلك الخيار. فالسؤال "ما هو وضعك في النظام المدرسي؟ "قد بعقبة بدائل مثل " إداري، مدرس، أمين مكتبة و "آخر" (حدد رجاءً)".

10- **اجعل الاستبيان موجزا قدر الإمكان بحيث يتطلب أقل وقت من المستجيب.** الأكثر احتمالاً هو أن المستجيبين يكملون الاستبيان القصير ويعيدونه. فيجب على الباحث أن يبذل جهدا في استبعاد جميع البنود غير الضرورية وخصوصا تلك التي تتوفر أجوبتها من مصادر أخرى. ويجب أن تخدم جميع بنود الاستبيان مهمة مشكلة البحث، أي يجب أن تستخلص البيانات المطلوبة لاختبار الفرضيات أو الإجابة على أسئلة الدراسة البحثية. مثلاً، السؤال الذي يسأل عن عمر المستجيب في دراسة حيث لا تكون هذه المعلومة مطلوبة في تحليل البيانات، يمكن استبعاده.

11- **تأكد أن لدى المستجيبين المعلومات الضرورية للإجابة على الأسئلة.** تجنب الأسئلة التي تتعامل مع خبرات أو موضوعات معروفة في كونها غير مألوفة لدى عينتك.

إجراء المقابلة

CONDUCTING THE INTERVIEW

سواء جرت المقابلة شخصياً أم هاتفياً فإن وظيفة المقابل الرئيسة هي أن يسأل سؤالا بطريقة يحصل فيها على استجابات صادقة ويسجل الاستجابات بشكل تام ودقيق.

المهمة الأولية للمقابل تكون بتوفير جو يضع المستجيب في وضع مريح. فبعد تقديم نفسه بطريقة ودية ينبغي على المقابل أن يوضح باختصار هدف المقابلة كما ينبغي أن يتجنب إعطاء معلومات كثيرة عن الدراسة مما قد يجعل المستجيب يتحيز. فمن المناسب البدء بالمقابلة بأسئلة بسيطة لا تبعث على الخوف.

ويتحمل المقابل أيضاً مسؤولية الحفاظ على اهتمام المستجيب مركزا على المهمة، وجعل المقابلة تمضي بصورة سلسة. ويمكن لهذا أن يجرى بشكل أفضل إذا كان المقابل يعرف بدقة الأسئلة وتسلسلها بحيث يمكنه طرح الأسئلة بصبغة دارجة دون توقف على نحو دائم لإيجاد السؤال الثاني. ويجب أن يمتنع المقابل عن إبداء الموافقة أو التعجب أو الصدمة لأي من إجابات المستجيب.

يمكن للمقابلة أن تكون منظمة تقريباً. ففي المقابلة الأقل تنظيماً، تطرح الأسئلة ذاتها على كل المستجيبين، لكن المقابلة تكون أكثر حواراً، وللمقابل حرية في ترتيب نظام الأسئلة أو إعادة صياغتها. وللحصول على بيانات جديرة بالمقارنة، يجب على المقابل، على أي حال، تقنين الإجراء باستخدام جدول لمقابلة منظمة. ويحتوي هذا الجدول على أسئلة محددة حسب نسق ثابت تُطرح على جميع المستجيبين، إضافة إلى فقرات واستقصاءات انتقالية. مثلاً، إذا بدأ المستجيب بالمراوغة، أو الاستطراد، أو تقديم ردود لا صلة لها، أو أساء تفسير السؤال، فعندئذ قد يستخدم المقابل أسئلة استقصائية ثابتة مثل "أوضح إجابتك أكثر" أو "هل بوسعك أن تقول لي أكثر من ذلك؟". وفي المقابلات الأقل تنظيماً، ينبغي توثيق أية انحرافات ملحوظة عن البروتوكول بحيث يمكن اخذ المعلومات بنظر الاعتبار لدى تحليل استجابة المستجيب. وباستخدام المجسات / الاستقصاء يجب على المقابل توخي الحذر في عدم الاقتراح أو التلميح بشأن الاستجابات المحتملة.

ويستغرق الأمر وقت تدريب أقل في تعليم المقابلين على إجراء مقابلة منظمة مما عليه الحال في المقابلة غير المنظمة، لأن كل شئ يحتاج المرء إلى قوله أو فعله موجود في جدول المقابلة. ولهذا السبب، فإن المقابلة المنظمة هي الأكثر شيوعاً في الاستخدام للدراسات الواسعة ذات المستجيبين الكثر.

تدريب المقابل Training The Interviewer

من الضروري بالنسبة للمقابلين المحتملين، الحصول على تدريب قبل ذهابهم للميدان. فنوعية المقابلين ربما هي إحدى أقل الجوانب تقديرا في البحوث المسحية. "فبالنسبة لعدد كبير من الأسئلة المسحية المألوفة، يمكن أن تؤدي مقابلات ضعيفي التدريب أو الإشراف إلى تقليص حجم العينة المؤثر إلى 20% أو 30% (Fowler, 1988, p. 148).

أما الموضوعات التي ينبغي تغطيتها في تدريب المقابل فلا بد أن تشمل ما يلي:

(1) إجراء الاتصال بالمستجيبين وتقديم الدراسة، (2) تعليمات حول طرح الأسئلة بحيث يطرح المقابلون جميع الأسئلة بطريقة مقننة ومتسقة، (3) إجراءات استقصاء الأجوبة غير الكافية بطريقة غير توجيهية، (4) إجراءات تسجيل الأجوبة على الأسئلة مغلقة النهاية ومفتوحة النهاية، (5) قواعد التعامل مع مظاهر علاقات التفاعل الاجتماعية للمقابلة بطريقة غير متحيزة (Fowler, 1988). ولغرض القدرة على الرد على أسئلة المستجيبين، ينبغي أن يعرف المقابلون كذلك هدف المشروع، ومن يرعاه، وطريقة المعاينة المستخدمة، والخطوات التي ستتخذ بشأن السرية.

وينبغي إعطاء المتدربين من المقابلين كتيبات مخطوطة حول إجراءات المقابلة. ويجب أن يشاهدوا المقابلات التي تجري عن طريق أشخاص مدربين، كما يجب الإشراف عليهم لدى إجراء تدريب على المقابلات. وفي هذا التدريب ينبغي أن يكون المستجيبون أشخاصا منتخبين من المجتمع نفسه الذي سيستخدم في المشروع البحثي.

استخدام الاستبيان البريدي
Using A Mailed Questionnaire

كما ناقشنا سابقاً، فإنه ليس بالأمر العملي دوما الحصول على بيانات مسحية تستخدم بنية المقابلة - في هذه الحالات تستخدم الاستبيانات، وهي في الغالب ترسل بريدياً للمستجيبين. ويعتبر الاستبيان جيد البناء عاملاً مهماً يؤثر على معدل الاستجابة. وقد قمنا سابقاً بتقديم إرشادات لكتابة الأسئلة. وفي هذا الجزء سوف ننظر في الترتيب الكلي للأسئلة، ورسالة الغلاف / التغطية والمتابعات، والعوامل الأخرى التي تسهم في إنجاح الاستبيان البريدي.

توجيهـــات Directions

من المهم البدء بتوجيهات دقيقة تبين للمستجيبين ما ينبغي فعله بالضبط. وضح كيف وأين ينبغي عليهم وضع إشارات استجاباتهم. مثلاً، "الرجاء إيضاح جوابك على الأسئلة التالية بوضع

علامة X في المربع المقابل لجوابك المختار "أو الرجاء استخدام قلم رصاص رقم 2 وأوضح إجاباتك بتظليل الدوائر المناسبة على ورقة الأجوبة الحسابية المنفصلة". فإذا تغيرت البنية داخل الاستبيان، فينبغي على المرء أن يعطي توجيهات جديدة لذلك الجزء.

نسق / نظام الأسئلة Order of Questions

حالما تكتب الأسئلة، فإنه ينبغي ترتيبها في نسق مناسب. ويعتبر نسق الأسئلة مهماً إذ بإمكانه التأثير على اهتمام المستجيبين في إكمال الاستبيان. وينبغي أن يكون السؤال الأول مثيراً للاهتمام بشكل خاص، وسهلاً بما يكفي لجميع المستجيبين في تفسيره والإجابة عليه. فإذا نشأت دافعية لدى المستجيبين للإجابة على السؤال الأول، فثمة احتمال أكبر في أن يستمروا مع الاستبيان. وينبغي أن يتحرى السؤال الأول عن معلومات ذات قيمة لها صلة واضحة بالموضوع قيد النظر. لهذا السبب ينبغي أن لا يبدأ الاستبيان مطلقاً بأسئلة لها صلة بالعمر، والجنس، والثقافة، والمهنة، والأصل العرقي، والحالة الزوجية وما إلى ذلك. فقد يعتبر المستجيبون أن هذه الأسئلة ليست ذات صلة أو كاقتحام للخصوصية، ولذلك قد يقررون عدم الاستمرار مع الاستبيان. وقد ينتهي الاستبيان في أقرب سلة مهملات. ويوصي أيضاً أن تكون الأسئلة القليلة الأولى ذات نهاية مغلقة، يمكن أن يكملها المستجيب بسرعة، بدلاً من الأسئلة مفتوحة النهائية التي قد تتطلب جواباً مكتوباً طويلاً.

يجب أن تُجمع الأسئلة المتشابهة في المحتوى معا. مثلاً، في استبيان يسأل أعضاء كلية معينة عن المهارات الأكاديمية الأساسية لطلابهم، يجب وضع جميع الأسئلة عن القراءة معا. ثم تظهر الأسئلة عن الكتابة معاً، يعقبها أسئلة لها علاقة بمهارات الرياضيات. وضمن مجالات المحتوى، يجب تجميع البنود حسب نوع السؤال. مثلاً، الأسئلة التي تتطلب ببساطة كلمة "نعم" أو "كلا" يمكن أن توضع معاً مثل البنود التي تتطلب من المستجيبين أن يرتبوا أو يوضحوا مدى الاتفاق من عدمه.

وفي كل مجال من مجالات الموضوعات يجب ترتيب الأسئلة حسب نظام نفسي مناسب. فالترتيب المنطقي أو النفسي يساهم في إجابات تم التفكير بها بشكل جيد من جانب المستجيبين. مثلاً، قد يود المرء أن يتيقن ما إذا كان المستجيبون راضين بظروف العمل قبل الطلب إليهم التوصية بالتغييرات. وإن طرحت أسئلة عامة وخاصة حول موضوع معين، فقم بوضع الأسئلة العامة أولاً. وينبغي أن تسبق البنود الموضوعية حول قضية أو وضع معين، الأسئلة الأكثر ذاتية. وينبغي للأسئلة التي يبدو الاعتراض عليها أقل احتمالاً، أن تسبق البنود التي يكون الاعتراض عليها أكبر. وقد يحجم بعض الناس أحياناً عن إجابة أسئلة حول الاتجاهات، والتفضيلات، والدوافع، والسلوك، والمشاعر الشخصية، وما شابه، لكنه إن أمكن استخدام أسئلة موضوعية أولاً لإيضاح وتحديد الموقف، فقد يكون من الأسهل للأشخاص تقديم الإجابة. مثلاً، الباحث الذي يود مسح الطلبة في إطار مدى تناولهم الماريجوانا، قد يبدأ بطرح أسئلة أكثر موضوعية

أولاً، مثل، "كيف تصف تناول المارجوانا في المدرسة: مشكلة خطيرة، مشكلة معتدلة، مشكلة زهيدة، أو لا مشكلة؟. ويمكن أن يتبع ذلك أسئلة مثل "هل تعتقد أن تكرار تدخين المارجوانا قد ازداد، أو بقي على حاله، أو تقلص هذا العام؟" و "هل تعرف الطلبة الذين يستخدمون المارجوانا؟" ثم ربما "هل دخنت المارجوانا؟".

أما مضمون المبدأ أعلاه في وضع أسئلة أقل عرضة للاعتراض قبل الأسئلة الأكثر عرضة للاعتراض، فهي أن الأسئلة المتعلقة بالبيانات الديموغرافية مثل العمر، والجنس، والمهنة ينبغي لها أن توضع في نهاية الاستبيان وليس في بدايته. وستكون هناك اعتراضات قليلة على إعطاء هذا النوع من المعلومات الشخصية بعد أن يكمل الشخص الاستبيان ويمكن أن يرى السبب وراء كون هذه الأنواع من البيانات ذات صلة.

معاينة المصفوفة Matrix Sampling

يستخدم أحيانا إجراء يدعى معاينة المصفوفة (Lord, 1962) عندما يكون المسح طويلاً ويكون المجتمع سهل المنال كبيراً. وينطوي هذا الأسلوب على اختيار المستجيبين عشوائيا حيث يعطي لكل واحد منهم مجموعة جزئية من الأسئلة التي تختار عشوائيا من مجموعة كل البنود. إن المزية العملية لاستخدام معاينة المصفوفة هي تقليص الوقت المطلوب لكل فرد في الإجابة. وهذه مزية مهمة لأن إحدى العقبات للحصول على معدل ردود عال هو عدم الرغبة من جانب بعض الأفراد في توفير وقت للرد على استبيان طويل. ويمكن للقارئ العودة إلى مرجع (Lord, 1962) للحصول على معلومات إضافية حول مفهوم معاينة المصفوفة.

بنية / تصميم الاستبيان Format Of The Questionnaire

ينبغي تصميم الاستبيان بحيث يكون جذاباً، وسهل القراءة والإجابة عليه، ومناسباً للباحث كي يعمل على ترميزه ووضع درجاته. ولتحقيق هذه الأهداف، ينبغي دمج الاقتراحات التالية في تصميم وثيقة المسح:

1- **رقّم الأسئلة على التعاقب خلال الاستبيان دون أي تكرار أو حذف.** فوضع عدد واحد لكل سؤال يجنبك الإرباك في ترميز الردود.

2- **يمكن التمييز بين الأسئلة عن أصناف الأجوبة وذلك باستخدام النمط المعتاد للأسئلة وأحرف كبيرة للأجوبة.** وتوضع أية توجيهات خاصة للإجابات داخل أقواس وتكتب بأحرف عادية(*). مثلاً:

(*) لاحظ أن هذا الاقتراح يخص اللغة الإنجليزية حسب الكتاب، وربما يمكن أخذ الاقتراح بعين الاعتبار لتنظيم الاستبيان باللغة العربية حسبما يكون مناسباً – (المراجع).

هل تفضل وضع معايير في المهارات الأساسية كشرط للتخرج من المدرسة؟ (حدد إجابتك).

______ 1) كلا

______ 2) نعم

3- **استخدم الأعداد لتحديد أصناف الإجابات المختلفة.** وتمثل الأرقام المخصصة للخيارات شكلاً من الترميز القبلي الذي سيسهل معالجة البيانات. ويمكن للمرء أن يضع فراغاً أو مربعاً أمام اختيارات الإجابة ويطلب من المستجيب أن يضع علامة (X) في الفراغ. مثلاً:

ما مستوى أعلى تعليم أكملته؟ (حدد أحدها).

______ 1) مدرسة ابتدائية.

______ 2) ثانوية ما.

______ 3) أكملت الثانوية.

______ 4) كلية ما.

______ 5) أكملت الكلية.

______ 6) عمل تخرج ما.

______ 7) درجة تخرج.

إن تحديد واحدة أو وضع (X) أمام البند (5) توضح أن الفرد قد أكمل الكلية، ويمكن بسهولة إجراء العدّ الخاص بالصنف (5). وبسبب الأطوال المختلفة للخيارات، فإنه يوصي بوضع خيارات الإجابات في البداية إلى يمين(*) أصناف الإجابات وليس إلى اليسار.

4- **حافظ على الثبات في تخصيص أعداد لأصناف الأجوبة المختلفة.** استخدم دائما العدد ذاته للإجابة ذاتها خلال كل الاستبيان. لقد جرت العادة على تخصيص أعداد متدنية للإجابات السالبة وأعداد عالية للإجابات الموجبة. مثلاً، يخصص العدد (1) إلى (كلا) ويخصص (2) إلى (نعم)، ويخصص (1) إلى (غير موافق)، (2) إلى (موافق). ومهما كان نظام الترقيم المختار، فيجب استخدامه بثبات، لأن من المربك للمستجيب ربط (1) مع (كلا) في الجزء الأول من الاستبيان ليجد أن (1) مرتبط مع (نعم) في جزء آخر.

5- **ينبغي ترتيب أصناف الإجابات بصورة رأسية وليس بصورة أفقية.** فالترتيب العمودي يجعل الاستبيان يبدو أقل ضغطاً ويقصي الخطأ المألوف في تحديد الفراغ في الجانب الخطأ من الجواب كما يمكن أن يحدث لما يأتي:

ما حالتك الزوجية الراهنة؟

(*) في الإنجليزية "يسار" . (المراجع)

_____ 1) لم أتزوج قط، _____ 2) متزوج، _____ 3) مطلق،
_____ 4) منفصل، _____ 5) أرمل.

إن الترتيب المربك أعلاه يمكن تحسينه بالطريقة التالية:

ما حالتك الزوجية الراهنة؟ (حدد واحدة)

_____ 1) لم أتزوج قط
_____ 2) متزوج
_____ 3) مطلق
_____ 4) منفصل
_____ 5) أرمل

6- **استخدم أسئلة موقوفة / معلقة (تتعلق بغيرها) عندما لا يكون كل سؤال مناسباً / ملائماً لجميع المستجيبين.** فالسؤال الموقوف / المعلق هو السؤال الذي تتوقف صلته على جواب في سؤال سابق. مثلاً، في مسح مصمم لتقدير اهتمام هيئة التدريس باستخدام الاختبار المعان بالحاسوب قد يستخدم السؤال التالي:

1. هل أنت مهتم باستخدام الاختبار المعان بالحاسوب في صفوفك؟

_____ 1) نعم
_____ 2) كلا (في حالة النفي الرجاء الانتقال إلى السؤال 6)
_____ 3) غير متأكد

إن كان الجواب بالإيجاب: هل تستخدم الاختبار المعان بالحاسوب في
_____ 1) الاختبارات القصيرة / السريعة (Quizzes)
_____ 2) الاختبارات
_____ 3) كلاهما

في المثال السابق يعتبر السؤال الثاني سؤالاً موقوفاً لأن صلته تتوقف على الجواب المقدم في السؤال الأول. فأعضاء هيئة التدريس الذين أوضحوا في السؤال (1) أنهم غير مهتمين بالاختبار المعان بالحاسوب، لن يتعين عليهم حتى قراءة الأسئلة الأربعة التالية، بل بوسعهم الانتقال إلى (6) وهو السؤال الملائم التالي. وهكذا، فإن الأسئلة الموقوفة / المعلّقة توفر الوقت للمستجيب وتقدم معلومات أكثر دقة للباحث. ويمكن إبراز الأسئلة الموقوفة بالأسهم أو بإشارات خاصة أو بشكل مغلق بخطوط مستقيمة (مربع، مستطيل).

7- **أعد نسخ الاستبيان بطريقة طباعة ذات نوعية عالية.** فالطباعة الممتازة تعطي الاستبيان مظهراً متميزاً أكثر وتعطي انطباعاً مفضلا أكثر لدى المستجيبين.

تسجيل الأجوبة Recording Answer

افترض النقاش السابق أن المستجيبين سيشيرون لإجاباتهم مباشرة على نموذج الاستبيان. ونوصي متى ما أمكن ذلك، أن يقدم الباحثون صحائف تدقيق ضوئي (Scannable sheets) ويوجهون المستجيبين على وضع إشارات لجميع الإجابات مباشرة على الصحائف. وتتوفر صحائف تدقيق ضوئي ذات أهداف عامة مع خمسة إلى عشرة خيارات لكل سؤال بكلفة رمزية. ويمكن قراءة الصحائف عن طريق موزع ضوئي (optical Scanner)، وتوضع البيانات على قرص وترسل مباشرة إلى الحاسوب لغرض تحليلها. إن استخدام صحائف التدقيق الضوئي يوفر الوقت ويقلص الخطأ البشري في ترميز الإجابات. ويبين الشكل 12.1 ورقة إجابة ذات عشرة خيارات. وتطبع أسئلة المسح على نموذج منفصل، وترفق صحيفة التدقيق الضوئي لغرض الإجابات. مثلاً:

1- كم عدد الفصول الدراسية للغة الأجنبية نفسها التي درستها في الثانوية؟

أ- لا يوجد (صفر)	د- ثلاثة	ز- ستة
ب- واحد	هـ- أربعة	ح- سبعة
ج- اثنان	و- خمسة	ط- ثمانية

ويمكن ترميز الإجابة على السؤال أعلاه مباشرة على صحيفة تدقيق الأجوبة.

وفي بعض الحالات يمكن طبع أسئلة المسح مباشرة على صحيفة تدقيق. ويبين الشكل 12.2 عينة من بعض الأسئلة المسحية المستخدمة في دراسة مفاهيم الطلبة الجدد عن كفاية مدارسهم الثانوية في إعدادهم للكلية.

الاختبار الميداني Field-Testing

قبل القيام بالطبع النهائي من الضروري أن يختبر الباحث الأداة بغية تحديد الغموض أو سوء الفهم أو أي نقص آخر. أولاً، من المناسب الطلب من الزملاء المطلعين على الدراسة اختبار مسودة الاستبيان وإعطاء آرائهم حول ما إذا كانت الأداة ستحوز على البيانات المطلوبة وما إذا كانوا يرون أية مشكلات صار تجاوزها سهواً.

وثانياً، يجب إجراء الاستبيان شخصيا وعلى نحو فردي على مجموعة صغيرة من الأفراد يتم اختيارهم من المجتمع الاحصائي الذي يجري اعتباره في هذه الدراسة. ويجيب المستجيبون على الأسئلة مرة واحدة ويقدمون تغذية راجعة للباحث حول أية صعوبات واجهوها في البنود. ويعطي اهتمام إلى تعليقات مثل " لا اعرف ما تعنيه هنا " و " أكثر من إجابة من هذه الإجابات تنطبق

الشكل 12.1: مثال على صحيفة تدقيق ضوئي لتسجيل الإجابات

GENERAL PURPOSE - NCS® - ANSWER SHEET

SEE IMPORTANT MARKING INSTRUCTIONS ON SIDE 2

SIDE 1

NAME (Last, First, M.I.)

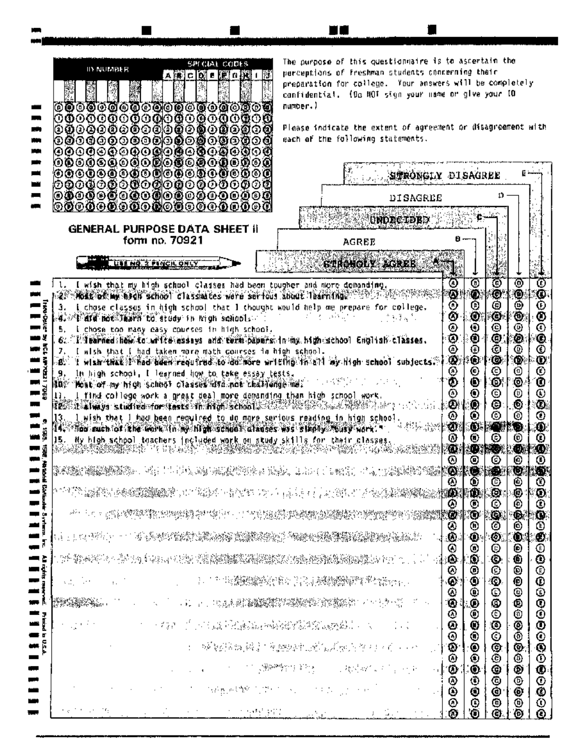
BIRTH DATE — MO. DAY YR.

IDENTIFICATION NUMBER — A B C D E F G H I J

SPECIAL CODES — K L M N O P

Source: Reproduced with permission from National Computer Systems, Inc.

الشكل 12.2: مثال على صحيفة تدقيق ضوئي مع عبارات مسحية


ID NUMBER

SPECIAL CODES

A B C D E F G H I J

The purpose of this questionnaire is to ascertain the perceptions of Freshman students concerning their preparation for college. Your answers will be completely confidential. (Do NOT sign your name or give your ID number.)

Please indicate the extent of agreement or disagreement with each of the following statements.

STRONGLY DISAGREE E

DISAGREE D

UNDECIDED C

AGREE B

STRONGLY AGREE A

GENERAL PURPOSE DATA SHEET II
form no. 70921

USE NO. 2 PENCIL ONLY

1. I wish that my high school classes had been tougher and more demanding.
2. Most of my high school classmates were serious about learning.
3. I chose classes in high school that I thought would help me prepare for college.
4. I did not learn to study in high school.
5. I chose too many easy courses in high school.
6. I learned how to write essays and term papers in my high school English classes.
7. I wish that I had taken more math courses in high school.
8. I wish that I had been required to do more writing in all my high school subjects.
9. In high school, I learned how to take essay tests.
10. Most of my high school classes did not challenge me.
11. I find college work a great deal more demanding than high school work.
12. I always studied for tests in high school.
13. I wish that I had been required to do more serious reading in high school.
14. Too much of the work in my high school classes was simply "busy work."
15. My high school teachers included work on study skills for their classes.

Source: Reproduced with permission from National Computer Systems, Inc.

علي". ويحاول الباحث أن يتبين ما إذا كان تفسير الأسئلة متماثلاً لدى جميع المستجيبين. وقد يوضح أحدهم بعض الأسئلة بطرق مختلفة لكي يرى ما إذا كانت ثمة أجوبة مختلفة قد أعطيت لصيغ متنوعة للسؤال ذاته.

إن الملاحظات التي تجري على المستجيبين أثناء ملئهم الاستبيان قد تنير البصيرة. فإنفاق وقت غير مستحق على سؤال أو ترك السؤال فارغا والعودة إليه فيما بعد يعطي إلماحاً / قرينة بأنه توجد هناك مشكلات في بعض هذه البنود.

ويمكن استخدام الاختبار الميداني لتوضيح البنود أو ربما حذف بعضها. ومن المهم بشكل خاص تحديد ما إذا كانت الأسئلة ستعمل على حد سواء بشكل جيد لدى الطبقات الاجتماعية المختلفة والجماعات الثقافية للمجتمع الإحصائي المدروس.

وتشمل بعض القضايا المحددة، التي ينبغي مراعاتها، في الاختبار الميداني ما يلي:

1- هل يبدو المستجيبون راضين عن الاستبيان ولديهم الدافعية لإكماله؟
2- هل بعض البنود مربكة؟
3- هل يمكن أن تؤدي بعض البنود إلى عداء أو إرباك من جانب المستجيبين؟
4- هل التعليمات واضحة؟
5- كم سيستغرق المستجيب من الوقت لإكمال الاستبيان؟
6- هل يفسر المستجيبون جميعهم البنود بذات الطريقة؟

إعداد رسالة الغلاف / التغطية Preparing The Cover Letter

قد يجد الباحثون أن من المفيد أن يبعثوا برسالة تمهيدية إلى المستجيبين المحتملين قبل إرسال الاستبيان نفسه. فهذا الإجراء ينبه الفرد للدراسة بحيث لا تكدره بشدة رزمة الاستبيان. وعلى أية حال، يجب أن تصاحب الاستبيان رسالة غلاف/ تغطية تخاطب المستجيب بالاسم والعنوان. ويبين الشكل 12.3 رسالة تغطية مع الأجزاء المهمة المحددة. وتقدم رسالة الغلاف المستجيبين المحتملين إلى الاستبيان و "تقنعهم" بالإجابة. وينبغي أن تشمل رسالة الغلاف العناصر التالية:

1- هدف الدراسة. ينبغي أن توضح الفقرة الأولى من الرسالة هدف الدراسة وفائدتها المحتملة. وقد يكون من المفيد ربط أهمية الدراسة بمجموعة مرجعية قد يتماثل الأفراد معها. مثلاً، ينبغي أن تؤكد رسالة الغلاف مع الاستبيان الموجهة إلى الطلبة الخريجين، على أهمية البيانات لأجل تحسين الدراسات العليا في الجامعة.

2- التماس التعاون. ينبغي أن توضح الرسالة سبب شمول المستجيب بالعينة وينبغي أن تنطوي على مناشدة التعاون من المستجيب. وينبغي أن يشعر المستجيبون بأن بوسعهم تقديم إسهام مهم للدراسة.

الشكل 12.3: مثال عن رسالة غلاف / تغطية لمسح معين

نص الرسالة	العنصر
جامعة إنديانا (.I.U)	ورقة العنوان ←
مكتب الدراسات التقويمية والاختبارات	
15/2/1994	التاريخ الأحدث ←
عزيزي خريج .I.U	
يجري مكتب الدراسات التقويمية والاختبارات مسحاً للخريجين الجدد من جامعة إنديانا لجمع معلومات عن اتجاهاتهم وآرائهم بشأن خبراتهم في الجامعة. إنا مهتمون بمدى تلبية .I.U لحاجاتك الأكاديمية. وسنستخدم نتائج المسح لمراجعة وتقوية البرامج للطلبة في الحاضر والمستقبل.	هدف المسح ←
لقد جرى اختيار اسمك في عينة عشوائية لجميع خريجي .I.U من 1988 حتى 1992. ولكي تمثل النتائج جميع الخريجين الجدد،	أهمية المستجيب ←
فمن المهم جداً إكمال الاستبيان وإعادته. ولا تتطلب الإجابة أكثر من عشر دقائق من وقتك، لكنها تعد حرجة لنجاح الدراسة. وأود أن أحثك على إكمال الاستبيان وإعادته في المغلف المرفق في	التماس التعاون ←
موعد لا يتجاوز 24/2/1994.	الوقت المحدد للإعادة ←
يمكنك الاطمئنان بأن إجابتك ستظل بسرية تامة. ولكل مغلف مستعاد عدد / رقم يمكننا من تدقيق الاسم في قائمة البريد لدى استعادة الاستبيان. وسيتم التخلص من المغلف بعدئذ. ولن يوضع اسمك مطلقاً على صحيفة الإجابة أو الاستبيان. وإذا كنت	ضمان الخصوصية ←
مهتماً في الحصول على خلاصة النتائج، فأرجو أن تحدد ذلك في المربع خلف المغلف، وسيتم إرسالها إليك قبل منتصف الصيف.	وعد بالنتائج ←
إذا كانت لديك استفسارات حول الدراسة، فأرجو أن تكتب أو تتصل تلفونياً بالرقم -----	
سيكون لتعاونك بالغ الامتنان	التعبير عن التقدير ←
مع كل إخلاص	
التوقيع	توقيع مدير المكتب ←
اسم المدير	

3- **الحماية المقدمة للمستجيب.** لا يجب أن تؤكد الرسالة للمستجيبين بأن إجاباتهم ستكون سرية فحسب بل يجب أيضاً أن توضح كيفية الحفاظ على السرية. ولغرض تسهيل إجراء المتابعة الضروري لمعدل إعادة مرتفع، فإنه يوصي باستخدام أعداد تمييز على الاستبيانات. فإذا لم يكن هناك تمييز، فإن مشكلة تحيز عدم الإجابة ستتعقد لأنه لا توجد هناك طريقة لمعرفة من أجاب ومن لم يجب كما تغدو إجراءات المتابعة مربكة. وإذا استخدمت أعداد التمييز، فمن الضروري إعلام المستجيبين بأن الأرقام موجودة لأنها تيسّر للباحث، تدقيق أسماء المستجيبين في قائمة البريد لدى إعادة الاستبيانات. ويجب التأكيد للمستجيبين بأن أسماءهم سوف لن توضع أبداً في الاستبيانات نفسها، وهكذا لن تكون هناك طريقة لربط إجابات معينة مع أي من الأفراد. وإذا أراد الباحث إتلاف الاستبيانات مباشرة بعد استعادة الإجابات، فإن هذه المعلومة ينبغي نقلها في الرسالة بغية إعادة التأكيد المستجيبين بإغفال أسمائهم.

ويفضل بعض الباحثين عدم استخدام نظام التعريف على الإطلاق، خصوصاً عندما يكون الموضوع حساسا. وفي هذه الحالة يكون من الضروري إدخال بطاقة بريدية في الرزمة البريدية يمكن أن يبعثها المستجيب بشكل منفصل لإيضاح أن الاستبيان قد تم إرساله أيضاً. وتحتوي البطاقة رسالة مطبوعة مسبقا بأن الاستبيان قد أعيد مع مجال للمستجيب لكتابة اسمه. وبهذه الطريقة يمكن الحفاظ على سجل للاستبيانات المعادة.

4- **الجهة الراعية للدراسة.** أن التوقيع على الرسالة مهم في التأثير على إعادة الاستبيان. فإذا كانت الرسالة جزءاً من أطروحة دكتوراه، فإنه من المفيد أن يقوم شخص معروف للمستجيبين، مثل رئيس قسم في الكلية أو عميدها بتوقيع الرسالة أو المصادقة عليها. إن مثل هذا التوقيع يمكن أن يكون أكثر تأثيرا من توقيع طالب خريج مجهول. وإذا كان هناك راع للدراسة مثل مؤسسة أو وكالة معينة فينبغي أن يذكر ذلك. وينبغي استخدام ورقة عنوان الجامعة أو الوكالة.

5- **الوعد بالنتائج.** قد يطرح عرض بالمشاركة في نتائج الدراسة مع المستجيبين إن كانوا مهتمين بذلك. وينبغي إعلامهم بكيفية تقديم طلبهم للنتائج المعروفة للباحث. وإحدى الطرق هي توفير مجال للتأشير / للتحديد على ظهر مغلف الإعادة ومجال لاسم وعنوان المستجيب كذلك.

6- **التقدير.** ينبغي شمول تعبير التقدير لمساعدتهم وتعاونهم في الدراسة.

7- **التاريخ الأحدث على الرسالة.** ينبغي وضع تاريخ لرسالة الغلاف قريباً من يوم إرسالها بالبريد. فالمستجيب المحتمل سوف لا يتأثر برسالة مؤرخة قبل عدة أسابيع من استلامها.

8- **التماس الإعادة المباشرة.** من المهم أيضا الحث على المباشرة للاستبيان. فإذا تم اقتراح مدة مقدارها أسبوعان أو شهر مثلاً، فإن المستجيب قد يركن الاستبيان جانباً وينساه رغم النوايا الطيبة. فالاستبيان الذي يخفق في كسب الانتباه خلال أسبوع لا يحتمل أن يعاد.

ينبغي شمول جميع العناصر أعلاه، غير أنه في الوقت ذاته، ينبغي أن تكون الرسالة مختصرة قدر الإمكان. وصفحة واحدة هي أقصى طول يوصي به. ضع الرسالة في مغلف مع الاستبيان. وضع دائما مغلفاً عليه عنوان ذاتي وعليه طابع للإعادة كي يستخدمه المستجيب. إن هذا لا غنى عنه من أجل معدل إعادة جيد.

يبين البحث أن نمط أجرة البريد المستخدمة قد يؤثر أيضاً على الإعادة. فالطوابع أوضحت أنها تسبب زيادة في معدل الإجابة أكثر من الرسالة المنتفخة ذات الأجر البريدي المطبوع عليها. ومن المؤكد أن الطابع يجعل الاستبيان يبدو شخصياً أكثر وأهميته أكبر، وأقل شبهاً ببريد سلة المهملات.

الحوافز المالية Monetary Incentive

أظهر مقدار كبير من البحوث أن استخدام الحافز المالي الرمزي يزيد من معدل الإجابة. فقد يكون المقدار المقدم هو مجرد مبلغ صغير لا يتجاوز الدولار أو ربع الدولار. ويعتقد أن النقود تخلق شعورا بالالتزام من جانب المتلقي والحاجة إلى الرد بالمقابل. وتكون الحوافز المالية أكثر تأثيراً عندما تدفع مسبقاً بدلا من الوعد بها، وعندما يتم تضمينها في البريد الأول للاستبيان وليس في المتابعة. وقد وجد (Zusman & Duby, 1984) أن تضمين حافز بدولار واحد قد حسن من معدل الإجابة بمقدار (19%) مقارنة مع المجموعة الضابطة التي لم تحصل على حافز. علاوة على ذلك، فقد وجد أن البريد الواحد مع الحافز المرفق طيا كان أكثر تأثيراً في استخلاص الإجابات من البريد والمتابعة لمجموعة مماثلة لم تحظ بحافز.

وبالطبع، فإن تقديم المال ليس ممكناً دائما لأنه، حتى المقدار الرمزي، قد يزيد من كلفة المسح بشكل كبير إن كانت العينة كبيرة. وعلى أية حال، فإنه خيار جدير بالاعتبار.

المتابعـــات Follow-Ups

تعد المتابعات المخططة للبريد أمراً جوهرياً للتوصل إلى أقصى نسبة من الإعادات في المسح الإستبياني البريدي. وتتخذ عادة عدة خطوات في المتابعات وهي موضحة هنا.

التذكير الأول First Reminder

إذا لم تتم إعادة الاستبيان في غضون أسبوع أو عشرة أيام بعد إرساله بالبريد ينبغي إرسال بطاقة بريدية إلى المستجيب. وهذه البطاقة تفيد كتذكرة أولية تشير إلى أن الاستبيان قد أرسل في وقت سابق وان الإجابة مهمة جدا للدراسة. قم بحث المستجيبين على إكمال و إعادة الاستبيان فوراً ("اليوم"). وبالطبع فإنه ينبغي إزجاء الشكر إلى أولئك الذين أعادوا الاستبيان.

ويمكن تقديم عرض بإرسال استبيان آخر إن كانت هناك حاجة له من قبل الذين أضاعوه أو لم يستلموه أبداً. وفي العادة فإن بطاقة التذكير البريدية ستعود بعدد كبير من الإجابات.

المتابعة الثانية Second Follow-Up

هذه المتابعة التي ينبغي إرسالها بعد ثلاثة أسابيع من الإرسال الأصلي تنطوي على رسالة ونسخة أخرى من الاستبيان ومغلف إعادة، عليه العنوان. وينبغي في الرسالة. أولاً، إعلام غير المستجيبين بأن استبياناتهم لم تستلم كما ينبغي تكرار فائدة الدراسة. وينبغي التأكيد على وجود الاستبيان البديل مع التماس قوي بإكماله وإعادته. ويجب إعلام المستجيبين بعدم الإجابة مرة ثانية إن كانوا قد أرسلوا الاستبيان بالبريد.

المتابعة الثالثة Third Follow-Up

ترسل المتابعة الثالثة والأخيرة بعد 6-7 أسابيع من البريد الأول. وهي تشبه المتابعة الثانية حيث فيها رسالة واستبيان بديل. ويبعث العديد من الباحثين هذه المتابعة بالبريد المسجل. فإذا توفر لدى الباحث 75%-90% من الاستبانات المعادة بعد ثلاث متابعات فإنه قد يكون مستعداً لإيقاف المسح واعتبار بقية الأفراد من غير المستجيبين. ويجب أن يقرر الباحث ما إذا كانت الإجابات التي تم الحصول عليها عبر نتائج المتابعة تستحق الكلفة والوقت المبذولين. ويشار أحيانا إلى أن يقوم الباحث في متابعته الثالثة بضم بطاقة بريدية يوضح عليها الأفراد بأنهم لا يرغبون في المشاركة بالمسح وسوف لن يعيدوا الاستبيان. أن مثل هذا الإجراء يسمح بتحديد دقيق لغير المستجيبين.

التعامل مع عدم الاستجابة Dealing With Nonresponse

يعتبر عدم الاستجابة مشكلة خطيرة في البحث المسحي. فما الذي سيفعله الباحث إزاء غير المستجيبين؟ كما أنه لا يسعه تجاهلهم ببساطة إن أريد للمسح أن يكون صادقاً. فاستخدام المعلومات من الذين يختارون الإجابة قد تتسبب في خطأ لأن المستجيبين يمثلون مجموعة مختارة ذاتياً قد لا تمثل أراء جميع العينة أو المجتمع الإحصائي.

فإذا ظل معدل الإجابة، بعد محاولات كل المتابعات، دون 75%، ينبغي على الباحث أن يحاول معرفة شئ عن صفات غير المستجيبين و/أو الحصول على إجاباتهم. وتبين البحوث أن المستجيبين يميلون إلى الاختلاف عن غير المستجيبين في صفات مثل الثقافة، والذكاء والدافعية، والاهتمام بموضوع المسح. فيمكن للمسح ذي المعدل المتدني للإجابة أن يكون متحيزا بشكل خطير، حتى وإن شرع الباحث بإرسال الاستبيان بريديا إلى عينة ممثلة. وينبغي على الباحث أن يحاول تحديد المدى الذي قد يختلف فيه المستجيبون عن غير المستجيبين. وثمة عدة طرق للقيام بذلك.

1- **قارن المستجيبين مع المجتمع الإحصائي**. إذا كان للباحث سبيل للوصول إلى المعلومات الخاصة بخصائص المجتمع الإحصائي - مثل العمر، والجنس، والثقافة، والحالة

الاجتماعية الاقتصادية، وما إلى ذلك - فبوسعه مقارنة خصائص المستجيبين مع خصائص المجتمع الإحصائي الذي انتخبت منه. فإذا كان المستجيبون بشكل عام نموذجاً للمجتمع الإحصائي في الخصائص المهمة، فبوسع الباحث افتراض أن المستجيبين يمثلون فعلاً مجموع المجتمع الإحصائي، وعليه يستطيع القيام بالتعميم من المستجيبين إلى العينة بأكملها. فإذا وجد أن المستجيبين مختلفون عن المجتمع الإحصائي، فيجب أن تقتصر النتائج على المستجيبين.

2- **قارن المستجيبين الأوائل بالمتأخرين.** لقد أوضحت البحوث أن غير المستجيبين مشابهون في الغالب للمستجيبين المتأخرين (Goldhor, 1974). وهكذا فإن الطريقة الثانية لتقدير إجابات غير المستجيبين هي إعادة الاهتمام بإجابات المستجيبين المتأخرين. وقبيل هذه الخطوة، ينبغي على الباحثين، على أية حال، تصنيف المستجيبين إلى مجموعتين أولى ومتأخرة ومقارنة إجاباتهم بهدف التحقق من وجود أية فروقات دالة. فإن لم توجد فروقات دالة بين المستجيبين الأوائل والمتأخرين، وجرى الاعتقاد بأن المستجيبين المتأخرين نموذج لغير المستجيبين، فحينئذ يكون بوسع الباحث الافتراض بأن المستجيبين يشكلون عينة غير متحيزة من المتلقين، وبهذا يمكن التعميم على كل المجموعة.

3- **قارن المستجيبين وغير المستجيبين.** ثمة طريقة منظمة، وتكون بإجراء مقابلة شخصية أو هاتفية مع عينة عشوائية صغيرة (ربما 10%) من غير المستجيبين. أن هذه العينة من غير المستجيبين المستمدة بهدف المقارنة تدعى أحيانا عينة الاختيار المضاعف " double-dipped sample" (Miller & Smith, 1983). وباستخدام الاستبيان كبرنامج مقابلة يجمع الباحث بعدئذ الإجابات من العينة العشوائية لغير المستجيبين. ويمكن المقارنة الإحصائية للوسط الحسابي للإجابات أو لنسبة إجابات غير المستجيبين مع إجابات المستجيبين لنرى ما إذا كانت المجموعتان تختلفان بشكل دال. فإذا لم توجد فروقات دالة عندما تقارن إجابات المستجيبين الأوائل مع إجابات عينة المقابلة فيمكن للباحث عندئذ أن يفترض بشكل معقول أن المستجيبين يمثلون عينة غير متحيزة لجميع الذين استلموا الاستبيان. ويمكن دمج البيانات وإجراء التعميمات على جميع العينة وعلى المجتمع الإحصائي. إلا أنه بدون مثل هذا التدقيق، ليس للمرء طريقة يعرف بها ما إذا كان المستجيبون مختلفين ومن ثم متحيزين. فاستخدام العينة ذات الاختيار المضاعف هي طريقة مفضلة لتدقيق الانحياز، إذ يمكن تقييم اتجاه ومدى الانحياز، بسبب عدم الإجابة، بصورة مباشرة. ومع ذلك فإنه أكثر كلفة واستهلاكاً للوقت.

ويبين (Aiken, 1981) أن المدى الذي تكون عنده إجابات المستجيبين على بنود المسح ممثلة لإجابات العينة بأكملها، هو دالة لحجم العينة، ونسبة الإعادة، ونسبة المستجيبين الذين أجابوا على البنود باتجاه محدد. وهو يقدم صيغة لتحديد أدنى نسبة من الناس الذين يجب أن يعيدوا المسح كي يشعر الباحث بالثقة بأن إجابات المستجيبين ممثلة لجميع العينة.

وإذا وجد المرء أن مجموعات فرعية محددة بشكل واضح لم ترجع الاستبيان، فقد يكون

من الضروري تغيير سؤال البحث الأصلي لاستبعاد هذه الجماعات الفرعية. فمثلاً، إذا أبدى مدرسو الثانوية معدل إعادة أدنى بكثير مما أبداه مدرسو الابتدائية في مسح معين، فقد يستنتج الباحث أن لدى مدرسي الثانوية صلة أو اهتماماً قليلين بالاستبيان فيقرر تحديد الدراسة بمدرسي الابتدائية. وتعاد صياغة سؤال البحث للإشارة إلى التغيير.

الصدق VALIDITY

يجب إعطاء الاهتمام بصدق المقابلات والاستبيانات - أي ما إذا كانت تقيس حقا ما يفترض قياسه. وينبغي أن يكون للمسح صدق ظاهري: إذ ينبغي أن يبدو صادقاً من أجل هدفه المقصود. فالأفراد يميلون أكثر إلى الإجابة على أسئلة يدركون صلتها ومعناها أكثر من أسئلة لا يستوعبون هدفها. ويكون الأفراد أقل ميلا إلى إكمال وإعادة استبيان يرونه غير مناسب.

النوع الأوضح لدليل الصدق العلمي، هو المرتبط بالمضمون / بالمحتوى، والذي قد يتم جمعه بمساعدة بعض الزملاء الكفوئين الذين يألفون هدف المسح، حيث يقومون بفحص البنود للحكم عما إذا كانت مناسبة لقياس ما يفترض قياسه، وما إذا كانت العينة ممثلة للمجال السلوكي قيد الدراسة.

لقد استخدمت بعض الدراسات الرصد المباشر للسلوك لتوفير دليل مرتبط بمعيار لصدق الإجابات. فبعد الحصول على الإجابات، يتم الرصد / الملاحظة لنرى ما إذا كان السلوك الفعلي للأفراد متفقا مع اتجاهاتهم وآرائهم وإجاباتهم التي عبروا عنها، أو أية إجابات أخرى. وقد تستخدم مصادر أخرى للبيانات، مثل أطراف ثالثة، كمعايير.

هناك متغيران مهمان يؤثران على صدق الاستبيان. أولاً، ما مدى أهمية الموضوع للمستجيب؟ فيمكننا افتراض إجابات صادقة أكثر من أشخاص مهتمين بالموضوع و/أو مطلعين عليه. ثانياً، هل يصون الاستبيان إغفال شخصية المستجيب؟ فمن المعقول، الافتراض انه سيتم الحصول على بيانات صدق أكبر إذا كان بوسع المستجيبين البقاء مجهولين، خصوصاً إذا تم طرح أسئلة حساسة أو شخصية.

الثبات RELIABILITY

يجب أن تتمتع بيانات المسح بالثبات لتحقيق الفائدة منها. فإذا لم تكن إجابات المستجيب متسقة / ثابتة، فإن صدق البحث موضع شك. واحد إجراءات تقدير ثبات الاستبيانات أو المقابلات هو وجود اثنين من المقابلين المختلفين ممن يقومون بمقابلة الأشخاص أنفسهم لتدقيق ثبات النتائج. ويمكن تدقيق الثبات الداخلي عن طريق بناء ما يزيد عن حاجة الأداة – بنود

حول الموضوع ذاته قد يعاد كتابتها وتكرارها في الاستبيان أو المقابلة وكلما كانت الإجابات متسقة كان الثبات عالياً.

ومن المحتمل تكرار الاستبيان أو المقابلة مع الأشخاص أنفسهم بعد فترة من الزمن أو تطبيق شكلين مختلفين من الاستبيان على ذات الأشخاص. إن مثل هذه الإجراءات باهظة في الغالب وتستغرق وقتا كثيراً مع ذلك، وهي غير عملية نوعا ما إذ ليس من السهولة إيجاد أفراد يرغبون في تكرار الاستبيان أو المقابلة. وثمة مشكلة أخرى مع هذه الطريقة، هي أن بعض الإجابات على الأسئلة التي تعالج مظاهر سلوكية أقل استقراراً، قد تتغير بشكل منطقي مع الزمن.

التحليل الإحصائي في المســـوحات
STATISTICAL ANALYSIS IN SURVEYS

لا تتطلب المسوحات عادة تحليلات إحصائية معقدة. فتحليل البيانات قد يتكون من تحديد التكرارات والنسب المئوية للإجابات على أسئلة الدراسة. مثلاً، قد يذكر مسح المصادر المكتبية عدد كتب الأعمال الخيالية وعدد كتب الأعمال الحقيقية، وما إلى ذلك. وإن مسح اتجاهات الناس حول قضية معينة، قد يذكر العدد والنسبة المئوية للمستجيبين الذين أعطوا كل إجابة مثل، "موافق بقوة"، "موافق"، "غير موافق" وما إلى ذلك.

الجدول 12.2 اتجاهات الطلبة نحو الزيادة في رسم النشاط

	موافق	بلا رأي	غير موافق	المجموع
الطلبة الجدد وطلبة الصف الثاني	60	68	32	160
طلبة السنتين الثالثة والأخيرة	80	46	66	192
الطلبة الخريجون	12	10	66	88
المجمـــوع	152	124	164	440

من المفيد تحويل الأعداد إلى نسب مئوية للقدرة على الحديث عن النسبة التي تجيب بطريقة معينة، وللقدرة على إجراء مقارنات. لننظر في بيانات التكرارات الافتراضية في الجدول 12.2 المستمدة من مسح (440) طالباً بشأن آرائهم حول الزيادة المقترحة لرسوم النشاطات التي ينتسب لها الطلبة في الجامعة. وبالنظر إلى التكرارات الخام قد يقول المرء إن طلاب الصفوف المتقدمة (66 في السنتين الثالثة والأخيرة) والطلبة الخريجين (66) يعارضون الزيادة في الرسوم.

لكن إن قام المرء بحساب "النسب المئوية" القائمة على أساس العدد الكلي للطلبة "في كل مجموعة" فبإمكانه أن يرى أن نسبة الطلبة الخريجين الذين يعارضون الزيادة مقارنة بطلبة السنتين الثالثة والأخيرة تزيد عن ضعفين. ويعزى ذلك إلى أن (66) من مجموع (88) من الطلبة الخريجين، أي (75%) لا يوافقون مقارنة مع (66) من (192) أو (34%) من طلبة السنتين الثالثة والأخيرة. وهكذا يبين الجدول أنه يحتمل أن لا يوافق الطلبة الخريجون على زيادة رسوم النشاطات بما هو أكثر من غير الخريجين.

ولتجنب سوء التفسير، ينبغي أن يتأكد المرء دائما بأن يعرض المجموع الكلي للمجموعات المختلفة في جداول مثل الجدول السابق 12.2.

إن الجداول التي تبين النسب المئوية لإجابات المجموعات المختلفة هي في الغالب أفضل طريقة لتوضيح العلاقة بين متغيرات المسح. وتدعى هذه الجداول بالجدولة المتقاطعة (Crosstabs) لأنها تسمح لأحدهم بالمقارنة عبر المجموعات. وتحوي أبسط الجدولة المتقاطعة متغيرين مع صنفين لكل متغير. والأشكال الأكثر تعقيداً ممكنة، على أية حال، مثل (2 × 2)، (4 × 2) ، (3 × 3) وما إلى ذلك. وتستخدم الجدولة المتقاطعة في الغالب مع أصناف أو بيانات اسمية.

ضبط المتغيرات في تحليل مسحي

Controlling Variables In A Survey Analysis

دعنا ننظر في النتائج الافتراضية لمسح خاص الاتجاهات نحو ضريبة مكتبية جديدة لتحسين وتوسيع المكتبة العامة للإقليم. ويبين الجدول (12.3) أن 63% (150/240) من سكان المدينة يفضلون ضريبة المكتبة مقارنة بـ 37% (150/240) من سكان الريف. ويبدو من هذه البيانات أن هناك علاقة بين مكان الإقامة والاتجاه نحو ضريبة المكتبة. وسيبين اختبار مربع كاي ما إذا كانت هناك علاقة دالة إحصائيا بين المتغيرين (انظر الفصل 6 لمناقشة مربع كاي). ويبين الجدول (12.4) حساب مربع كاي لهذه البيانات. والتكرارات المتوقعة لكل خلية تتضح بين الأقواس.

الجدول 12.3: الاتجاهات نحو ضريبة المكتبة حسب الإقامة

	مدينة	ريف	المجموع
موافق	150	96	246
معارض	90	164	254
المجموع	240	280	500

الجدول 12.4: التكرارات الملاحظة والمتوقعة لبيانات الاتجاهات

	مدينة	ريف	المجموع
موافق	150(118)	96(128)	246
معارض	90(122)	164(132)	254
المجموع	240	260	500

$$X^2 = \frac{(150-118)^2}{118} + \frac{(96-128)^2}{128} + \frac{(90-122)^2}{122} + \frac{(164-132)^2}{132}$$

$$X^2 = 32.83$$

بالعودة إلى الجدول (A.4) مع درجة حرية واحدة، يمكن للمرء أن يرى أن مربع كاي وهو (32.83) دال جداً (01. >). ونستنتج أن هناك علاقة دالة إحصائية بين مكان الإقامة والاتجاه نحو ضريبة المكتبة. وقد يشير المراقب الأكثر حذراً، على أية حال، إلى أن المدينة (وهي موقع الجامعة الرسمية الكبرى) لها نسبة أعلى من الناس المثقفين / المتعلمين وأن المستوى الثقافي / التعليمي وليس مكان الإقامة، بحد ذاته، هو الذي يعزى إلى الاتجاه المفصل نحو ضريبة المكتبة.

وبهدف استطلاع هذا التفسير البديل، يمكننا ضبط متغير المستوى الثقافي وذلك بجعله ثابتاً، ثم نلاحظ ما إذا كانت العلاقة بين المتغيرين الأولين تظل قائمة. إن أبسط طريقة تجعل المتغير ثابتا هي بتقسيم الأفراد إلى مجموعات منفصلة، لكل منها قيمة مختلفة في ذلك المتغير، ثم النظر إلى الجدولة المتقاطعة لكل من هذه المجموعات بشكل منفصل. في هذه الحالة يمكننا تصنيف (500) مستجيب إلى ذوي ثقافة جامعية وبدون ثقافة جامعية (بافتراض تيسر هذه المعلومات) والنظر في العلاقة داخل المجموعتين المنفصلتين.

الجدول 12.5: الاتجاهات نحو ضريبة المكتبة المتعلقة بالتعليم / الثقافة وليس الإقامة

	ثقافة جامعية		
	مدينة	ريف	المجموع
موافق	144	80	224
معارض	36	20	56
المجموع	180	100	280
	ثقافة غير جامعية		
	مدينة	ريف	المجموع
موافق	6	16	22
معارض	54	144	198
المجموع	60	160	220

توضح البيانات في الجدول (12.5) أن التفسير البديل صحيح. فإذا تم ضبط الثقافة الجامعية بجعلها ثابتة، فإنه لا توجد علاقة بين المتغيرين (مكان الإقامة) و(الاتجاه نحو الضريبة). فبين ذوي الثقافة الجامعية يفضل 80% (144/180) من سكان المدينة هذه الضريبة، ويفضل ذلك 80% (80/100) من سكان الريف. ومن بين الذين ليس لديهم ثقافة جامعية (6/60) 10% من سكان المدينة يفضلون هذه الضريبة، وكذلك يفضلها 10% (16/160) من سكان الريف.

لنفترض أن البيانات كانت كما في الجدول (12.6). فمع توفر ثبات المستوى الثقافي داخل كل جدول، فإن العلاقة بين المتغيرين (الإقامة) و (الاتجاه) واضحة. فمن بين ذوي الثقافة الجامعية يفضل الضريبة 67% (80/120) من سكان المدينة مقارنة مع 33% (20/60) من سكان الريف. ومن بين الذين ليس لديهم ثقافة جامعية، فإن 58% من سكان المدينة يفضلون الضريبة مقارنة مع 38% من سكان الريف. وفي هذه الحالة، هناك شيء ما حول المستجيبين في المدينة والريف، عدا ثقافتهم، هو الذي يقودهم إلى الإجابة بشكل مختلف حول ضريبة المكتبة.

قد يود المرء كذلك التحقق من فروقات الجنس في الإجابات في مسح معين. في هذه الحالة ستبين الجدولة المتقاطعة تكرار الإجابات على الأسئلة لكل من الذكور والإناث بشكل منفصل. ويمكن فحص فروقات الطبقة الاجتماعية بتصنيف الأفراد في مجموعات منفصلة على أساس المعايير المناسبة والنظر في إجابات كل مجموعة.

الجدول 12.6: الاتجاهات نحو ضريبة المكتبة حسب الإقامة وليس الثقافة

	ثقافة جامعية		
	مدينة	ريف	المجموع
موافق	80	20	100
معارض	40	40	80
المجموع	120	60	180
	ثقافة غير جامعية		
	مدينة	ريف	المجموع
موافق	70	76	146
معارض	50	124	174
المجموع	120	200	320

إحصاء الجدولة المتقاطعة Statistics For Crosstabs

تستخدم الجدولة المتقاطعة بشكل واسع لتوضيح الفروقات بشكل بياني في الإجابات بين المجموعات المختلفة، وما إذا كانت العلاقة موجودة أم غير موجودة بين المتغيرات. وحين يقاس كلا المتغيرين في الجدولة المتقاطعة حسب مستوى اسمي، فقد يستخدم اختبار مربع كاي لتحديد ما إذا كانت هناك علاقة منتظمة موجودة بين المتغيرين. وعلى أية حال، سيوضح مربع كاي "فقط" ما إذا كانت المتغيرات مرتبطة أو مستقلة. إنها لا تدل على مدى ارتباطها. مثلاً، أن قيمة (X^2) لبيانات الثقافة الجامعية في الجدول 12.6 هي (17.09) التي تعد دالة إحصائياً عند مستوى (.01). وهكذا نعرف أن هناك علاقة دالة بين مكان الإقامة والاتجاه نحو الضريبة بين المثقفين جامعياً، لكننا لا نعرف مدى قوة ارتباط هذين المتغيرين.

ولمعرفة مدى العلاقة، يجب على المرء حساب معامل الارتباط. ومعامل الارتباط الذي يستخدم على نحو متكرر مع البيانات الاسمية في جداول (2 × 2) هو معامل الارتباط فاي (ϕ). والمعامل ϕ تبسيط رياضي لمعامل بيرسون التتابعي جداول (2×2). وهكذا فإن لفاي القيمة (صفر) عندما لا توجد هناك علاقة، فهي (+1.00) في حالة العلاقة الموجبة التامة و (−1.00) في حالة العلاقة السالبة التامة. أن معامل فاي لذوي الثقافة الجامعية في الجدول (12.5) هو (.32). ونفسر معامل كاي مثل أي معامل بيرسون أخر. إن (.32) تبين علاقة موجبة ضعيفة نسبياً بين مكان الإقامة والاتجاه نحو ضريبة المكتبة بين ذوي الثقافة الجامعية.

بالنسبة للجداول التي هي أكبر من (2×2) فإن قياساً مناسباً للعلاقة هو الإحصاء كابا (K). فإذا كانت هناك علاقة تامة بين المتغيرات فسيكون (K) مساويا (1.00). وإذا كان الاتفاق بين المتغيرات هو تماما ما هو متوقع من خلال الصدفة فإن (K) يساوي صفراً. وإذا كان الاتفاق أقل مما هو متوقع بالصدفة فإن (K) سيكون عدداً سالباً.

وإذا كان كلا المتغيرين في الجدولة المتقاطعة ترتيبيين فإن إحصاءات مثل معامل ارتباط كندال للتوافق (W) أو إحصاءة غاما يمكن استخدامها لإيضاح قوة العلاقة بين المتغيرات. أن مناقشة شاملة لإجراءات الارتباط للمتغيرات الاسمية و الترتيبية يمكن إيجادها في (Siegel & Castellan, 1988).

الخلاصة SUMMARY

يستخدم المسح بشكل واسع كطريقة بحث لجمع البيانات التي تتراوح عدّ/ إحصاء مادي وتكرارات إلى اتجاهات وآراء. وتصنف المسوحات وفق التركيز والمدى ووقت جمع البيانات. وينبغي أن تنطوي على تخطيط دقيق، ومعاينة غير متحيزة للمجتمع الإحصائي، وتطوير مدروس لأدوات جمع البيانات وتحليل دقيق للنتائج.

وإذا استخدم الباحثون نوعا معينا من المعاينة الاحتمالية، فيمكنهم أن يستنتجوا قيم المجتمع الإحصائي من نتائج العينة. والإجراء الاعتيادي هو إقامة فترة ثقة يكون الأكثر احتمالاً لاحتواء قيمة المجتمع الإحصائي. أما طول / مدى الفترة فهو دالة للمجازفة التي يرغبون في اتخاذها بشأن وقوعهم في الخطأ وحجم العينة. وتضيق الفترة عندما يتقلص احتمال الصواب وعندما يزداد حجم العينة. وبوسع المرء كذلك استخدام هذا الإجراء لتقدير حجم العينة المطلوبة إلى أي مستوى مرغوب من الدقة.

المقابلات والاستبيانات هما الأداتان الرئيستان لجمع البيانات من أجل المسح. ويشمل كلا الإجراءين طرح أسئلة عن موضوعات مختارة، لكن لكل منها مزايا ومساوئ فريدة. من المهم أن تكون الأدوات المستخدمة ذات صدق وثبات. وبرهنت إجراءات المتابعة المتنوعة عن فعاليتها في زيادة الإعادات من الاستبيانات البريدية.

تؤمن الجدولة المتقاطعة طريقة ممتازة لإبراز العلاقة الموجودة بين المتغيرات في مسح معين.

مفاهيم أساسية Key Concepts

census	إحصاء/ تعداد السكان
checklists	قوائم تدقيق / مراجعة
closed-ended question	سؤال مغلق النهاية
cohort study	دراسة الجماعة
contingency question	سؤال موقوف / معلّق
cross-sectional survey	مسح عرضي
crosstabs	جدولة متقاطعة
double-barreled question	سؤال مزدوج
field-testing	اختبار ميداني
interview	مقابلة
interviewer bias	تحيز المقابل
Likert-type item	بند من نمط ليكرت
longitudinal survey	مسح طولي
margin of error	هامش الخطأ
nonresponse	عدم الاستجابة
open-ended question	سؤال مفتوح النهاية
panel study	دراسة مجموعة بذاتها
probe	تفحص دقيق

ranking items	بنود ترتيب
response set	مجموعة استجابات
sample survey	مسح العينة
scaled items	بنود مقيسة
social desirability bias	تحيز الرغبة الاجتماعية
trend study	دراسة النزعة / الاتجاه

تمارين EXERCISES

1- اقترح سؤالاً بحثياً يكون المسح أفضل إجابة عليه.

2- ما أسلوب جمع البيانات الذي تستخدمه لكل مما يلي من المسوحات؟

أ. مسح عينة من مدرسي الابتدائية على امتداد الدولة يخص أساليب الانضباط المستخدمة في مدارسهم.

ب. مسح آراء الناس في عاصمة / حاضرة كبيرة حول طريقة المعالجة الراهنة للأحداث الذين يرتكبون جرائم عنف، في النظام القضائي للدولة / للولاية.

ج. مسح لبعض الخصائص غير المعرفية لصف من الطلبة الجدد وعلاقتها بالتحصيل الأكاديمي خلال السنة الأولى من الكلية.

د. مسح لـــ 500 من الناس على امتداد الولايات المتحدة حول ما إذا كانوا يساندون مقترحات الرئيس حول الرعاية الصحية.

3- كيف تؤثر نسبة الردود نعم / كلا على حجم العينة المطلوبة؟

4- ذكرت أخبار المساء التلفزيونية ما يلي: في مسح أجري مؤخرا على الشعب الأمريكي، قال 45% من المستجيبين أنهم موافقون على أداء الرئيس. ما الذي تود أن تعرفه قبل قيامك بتفسير هذا التقرير؟.

5- كيف تفسر التقرير التالي؟ وجد استفتاء (1000) ناخب مسجل اختيروا بشكل عشوائي في ولاية أنديانا، أن 37% يفضلون اليانصيب الرسمي. إن الأعداد من هذا الاستفتاء عرضة لخطأ معاينة بمقدار ±3%. مستوى الثقة هو 95%.

6- لكل من العينات الثلاث أدناه، أنشئت فترة الثقة 0.95 لنسبة المجتمع الإحصائي. افترض أن نسبة العينة هي (.40) لكل منها. ما هو تأثير أي زيادة في حجم العينة على طول الفترة؟ لماذا؟

عدد أفراد العينات: أ- 100 ، ب- 1000 ، ج- 10,000

7- قمت بتطوير مقياس لقياس استنـــزاف طاقة المدرسين. وكان لعينة عشوائية مقدارها (100) مدرس يعملون في نظام مدرسي لمدينة كبرى / للعاصمة متوسط درجات (10.5) مع انحراف معياري (2.3). ما تقديرك لمتوسط درجات استنـــزاف الطاقة لكل مجتمع المدرسين في النظام المدرسي؟ استخدم مستوى الثقة 95%.

8- تود مؤسسة استفتاء قومية أن تكون قادرة على التنبؤ بنتيجة انتخابات الرئاسة ضمن ±5%. ما مدى حجم العينة العشوائية المطلوبة لتحقيق هذا المستوى من الدقة؟ افترض مستوى ثقة 95%.

9- كان لمسح معين معدل إجابات أولي مقداره 51%. ما هي المقترحات التي ستقدمها للباحث للتعامل مع معدل الإجابة المتدني هذا؟

10- الاهتمام باختصاص رئيسي في مجال الأعمال قد اخذ بالتدني في الجامعة خلال السنوات القليلة الماضية. ويريد قسم إدارة الأعمال أن يعرف ما إذا كان لدى صف الجدد أي اهتمام بهذا الاختصاص. ومع عدم وجود الموارد المالية لمسح جميع الطلبة الجدد (4500) طالب جديد فالهم قاموا بمسح عينة عشوائية من (500). وقد وجدوا أن (110) من الطلبة ذكروا أنهم مهتمون بمثل هذا الاختصاص. وعلى مستوى الثقة (95%)، ما هو تقديرك لعدد الطلبة الجدد الذين قد يهتمون بهذا الاختصاص في قسم إدارة الأعمال.

11- أي مما يلي تعتبر عينات متحيزة من مجتمع طلبة الكليات في جامعة كبيرة؟

أ. عينة عشوائية من الطلبة الذين يدخلون إلى المكتبة مساء الجمعة.

ب. عينة عشوائية من الطلبة المسجلين في الصفوف.

ج. عينة عشوائية من الطلبة الذين يشترون بطاقات موسمية لمباريات كرة السلة.

د. عينة مكونة من طلبة تطوعوا بعد مشاهدة إعلان في جريدة الكلية.

12- افترض انك تجري مسحا لتحديد الكيفية التي يشعر بها مدرسو مدارس ابتدائية في إقليم حول سياسة حجز التلاميذ.

أ. اكتب سؤالين مغلقي النهاية لهذا المسح.

ب. اكتب سؤالين مفتوحي النهاية لهذا المسح.

ج. اكتب سؤال موقوفاً / معلّقاً.

13- بالاعتماد على وقت جمع البيانات، صنف كلاً من المسوحات التالية:

أ. دراسة تيرمان عن الراشدين الذين كانوا موهوبين كأطفال.

ب. مقارنة تحصيل الرياضيات في المدارس الحكومية في 1974، 1984، 1994.

ج. متابعة خريجي عام 1990 من كلية إدارة الأعمال بجامعة انديانا.

د. مسح التحصيل القرائي في مستويات السنوات الابتدائية المختلفة في نظام مدرسي معين عام 1994.

14- أي مما يلي يعد مزية لنمط المسح الطولي.

أ. دراسة فردية أكثر تركيزاً.

ب. توفير بيانات لمجموعات عمريه مختلفة في الوقت ذاته.

ج. جمع بيانات فورية.

د. عدم وجود أخطاء معاينة.

15- تضمنت مجلة أخبار شعبية أسبوعية استبياناً مفصلاً حول الاتجاهات إزاء الجريمة. ودعا المحررون القراء إلى ملء الاستبيان و إعادته إلى المجلة. ونشرت المجلة عقب ذلك مقالة حول الخوف الواسع من الجريمة بين المواطنين الأمريكان، وخصوصاً كبار السن. كيف تقيّم هذا المسح؟

الأجوبة ANSWERS

1- الأجوبة ستختلف.

2- أ. استبيان بريدي،

ب. مقابلة هاتفية،

ج. استبيان يجرى تسليمه بشكل مباشر،

د. مقابلة هاتفية.

3- كلما كانت النسب قريبة من ‎.50-.50‎، كان حجم العينة المطلوب كبيراً.

4- سوف يود المرء معرفة حجم العينة، وكيف استمدت، وطول الفترة حول معْلمة (parameter) التقديرية للمجتمع (هامش الخطأ) ومستوى الثقة.

5- إن فترة الثقة البالغة (‎.95‎) لنسبة الناخبين المسجلين ممن يفضلون اليانصيب الحكومي هي ما بين 34، 40 في المئة.

6-

العينة ج	العينة ب	العينة أ
$= \sqrt{\frac{(.40)(.60)}{100}}$	$= \sqrt{\frac{(.40)(.60)}{100}}$	الخطأ المعياري $= \sqrt{\frac{(.40)(.60)}{100}}$
$= .005$	$= .0155$	$= .049$
$= .40 \pm 1.96\ (.049)$	$= .40 \pm 1.96\ (.0155)$	الفترة = $.40 \pm 1.96\ (.049)$
$= .40 \pm .01$	$= .40 \pm .03$	$= .40 \pm .10$
= .39 إلى .41	= .37 إلى .43	= .30 إلى .50

إن زيادة في حجم العينة يقلص طول الفترة. وعندما تزداد (n) يقل خطأ العينة ومن ثم حجم الفترة.

7- الانحراف المعياري $= \frac{s}{\sqrt{n}} = \frac{2.3}{\sqrt{100}} = \frac{2.3}{10} = .23$

التقدير $= \overline{X} \pm 1.96(.23) = 10.5 \pm .45$

أو بين 10.05 و 10.95

8- افترض $p = q = .50$

$$1.96\sqrt{\frac{(.50)(.50)}{n}} = .05$$

$$\sqrt{\frac{.25}{n}} = \frac{.05}{1.96}$$

$$\frac{.25}{n} = \left(\frac{.05}{1.96}\right)^2$$

$$\frac{.25}{n} = 0.00065$$

$$n = 384.6$$

أو $n = 385$

9- ينبغي على الباحث المتابعة ببطاقة تذكير بريدية ثم إرسال بريد آخر أو اثنين من الاستبيان. وبعد استكمال محاولات المتابعة ينبغي أن يحاول الباحث مقابلة بعض الباقين من غير المستجيبين لمعرفة خصائصهم والحصول على إجاباتهم بغية تحديد ما إذا كانوا يختلفون عن المستجيبين بصورة دالة.

10- بين 18% و 26% أو بين 810 و 1170 من الطلبة يمكن توقع اهتمامهم بالاختصاص الرئيسي في قسم إدارة الأعمال كما هو محسوب هنا:

$$p = .22\left(\frac{110}{500}\right)$$

$$q = .78$$

(الخطأ المعياري) $s.e. = \sqrt{\frac{(.22)(.78)}{500}} = .0185$

الفترة = $.22 \pm 1.96(.0185) = .22 \pm .04$

11- العينات أ، ج، د لا تكون ممثلة لمجتمع طلبة الكلية في جامعة كبيرة.

12- ستختلف الإجابات.

13- أ. دراسة طولية لمجموعة بذاتها ،
ب. دراسة طولية للنــزعة ،
ج. دراسة طولية لجماعة ،
د. مسح عرْضي .

14- أ

15- إن الذين اكملوا وأعادوا الاستبيان ليسوا عينة ممثلة لجميع المواطنين الأمريكان. وثمة عدد من العوامل ستعمل على تحيز هذه العينة كالمستوى الاجتماعي الاقتصادي، والمستوى الثقافي، والاهتمام الكافي بالموضوع لإكمال الاستبيان، وأجرة البريد للإعادة.

REFERENCES المصادر

Aiken, L.R. (1981). Proportion of returns in survey research. *Educational and Psychological Measurement,* 41, 1033-1038.

Blonston, G, (1994, February 1). Drug abuse by young teens shows ominous rise. *Indianapolis Star;* p. A1.

Cage, M.C. (1994, January 26). Beyond the B.A. *Chronicle of Higher Education,* 40(21), A29.

Coleman, J.S., et al. (1966). *Equality of educational opportunity.* Washing- ton, DC: U.S. Government Printing Office.

Coughlin, E.K. (1990, February 7). Researchers practice the science and art of public-opinion polling. *Chronicle of Higher Education,* 36, A6, A10-11.

Dillman, D.A. (1978). *Mail and telephone surveys.* New York: John Wiley.

Emery, E.M., Ritter-Randolph, G.P., Strozier, A.L., and McDermott, R.J. (1993). Using focus group interviews to identify salient issues concerning college students' alcohol abuse. *Journal of American College Health,* 41, 195-198.

Fowler, F.J. (1988). *Survey research methods.* Newbury Park, CA: Sage Publications;

Goldhor, H. (1974), *The use of late respondents to estimate the nature of non-respondents.* Washington, DC: U.S. Office of Education. (ERIC Document ED 083309)

Jacobs, L.C. (1985). *College freshmen view their high school preparation.* Indiana University: Indiana Studies in Higher Education, No.53.

Lord, F.M. (1962). Estimating norms by item-sampling. *Educational and Psychological Measurement,* 22, 259-267.

Lord, F.M., and Novick, M.R. (1968) .*Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Miller, L.E., and Smith, K.L. (1983). Handling nonresponse issues. *Journal of Extension,* 21, 45-50.

Mitchell, M., and Jolley, J. (1988). *Research design explained.* New York: Holt, Rinehart and Winston.

Morgan, D.L. (1988). *Focus groups as qualitative research.* Beverly Hills, CA: Sage Publications.

Siegel, S., and Castellan, N.S. (1988). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

Smith, T. W. (1987). That which we call welfare by any other name would smell sweeter: An analysis of the impact of question wording on response patterns. *Public Opinion Quarterly,* 51, 75-83.

Stewart, D. W., and Shamdasani, P.N. (1990). *Focus group theory and practice.* Newbury Park, CA: Sage Publications.

This year's freshmen: A statistical profile. (1994, January 26). *Chronicle of Higher Education,* 40(21), A30-31.

Wilhoit, G.C., and Weaver, D.H. (1990). *Newsroom guide to polls & surveys.* Bloomington: Indiana University Press.

Zuhl, H. (1994, February 4). Ongoing IU study checking back with smokers after 13 years. *IU Newspapel;* 18, p. 12.

Zusman, B.J., and Duby, P.B. (1984). An evaluation of the use of token monetary incentives in enhancing the utility of post-secondary survey research techniques. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.



البحث النوعي والتاريخي
Qualitative and Historical Research

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل، سيكون بوسع الطالب أن:

1- يميز بين البحثين الكمي والنوعي.
2- يصف طرق الباحثين النوعيين لإقامة الثقة.
3- يصف الجوانب الحرجة للبحث النوعي.
4- يصف طبيعة دراسات الحالة ومزاياها وقصورها.
5- يميز بين دراسات الحالة والدراسات الطبيعية، وتجارب الموضوع الواحد.
6- يميز بين رصد المشارك وغير المشارك.
7- يدرج أنماط الرصد/ الملاحظة لغير المشاركين.
8- يعرف تحليل المحتوى / المضمون ويوضح هدفه.
9- يعرّف علم وصف الأعراق/ الأجناس (الأثنوجرافيا) ويعطي مثالاً على دراسة أثنوجرافية.
10- يصف خصائص البحث التاريخي.
11- يميز بين المصادر الأولية والثانوية في البحث التاريخي.

12- يعرّف ويعطي أمثلة للنقد الداخلي والخارجي.

إن طرق البحث الموصوفة في الفصل التاسع حتى الثاني عشر تستخدم الأعداد للإجابة عن الأسئلة. وتصف مثل هذه الإجراءات بكونها بحوثاً كمية لأنها تستخدم قياسات كمية مثل التكرارات، والأوساط الحسابية، والارتباطات، والاختبارات التائية. وخلافاً لذلك، تستخدم البحوث النوعية كلمات للإجابة على الأسئلة. أما البحث التاريخي فإنه، على الأغلب، نوعي في طبيعته رغم أن الطرق الكمية تستخدم عندما تكون مناسبة.

البحث/ الاستقصاء النوعي QUALITATIVE INQUIRY

إن العبارة "بحث نوعي" تشير إلى مصطلح شامل لأنواع مختلفة من أساليب البحث التربوي والتقييم حيث توسم بأسماء مختلفة مثل الأثنوجرافيا، والبحث الطبيعي، ودراسات الحالة، والعمل الميداني، والدراسات الميدانية، ورصد المشارك. ويمكن التمييز بين هذه الطرق في إطار التقاليد الفلسفية والتحليلية المختلفة. ومع ذلك، فإنها تشترك في جملة مظاهر مألوفة تفصلها عن الأسلوب الكمي في البحث التربوي والتقييم الذي تجد له وصفاً عاماً في الفصول 1، 2، 4 والإجراءات الموصوفة في الفصل 9 حتى الفصل 12. وننصح القارئ المهتم بإجراء دراسة لبحث نوعي بمراجعة النصوص حول الطرق النوعية مثل (Bogdan & Biklen, 1992) أو (Lancy, 1993).

التمييز بين البحث النوعي والبحث الكمي

Distinguishing Qualitative Inquiry from Quantitative Inquiry

يختلف البحث النوعي عن الأسلوب الكمي في دراسة الظواهر الاجتماعية والسلوكية في رفضه للحجة التي تقول بأن هدف وطرق العلوم الاجتماعية هي، من ناحية مبدئية على الأقل، الهدف والطرق ذاتها الخاصة بالعلوم الطبيعية أو المادية(*). فالباحثون في النمط الكمي يناقشون بأن كلاً من العلوم الاجتماعية والعلوم الطبيعية تناضل من أجل نظريات قابلة للاختبار والتأكيد / التثبيت، لتفسير الظواهر عن طريق توضيح كيفية اشتقاقها من افتراضات نظرية (أنظر النقاش الخاص بالنظرية العلمية في الفصل 1). وبعبارة أخرى، يهدف كلاهما إلى نوع من التفسير

(*) إن وجهة النظر التي تقول بأنه ينبغي أن تكون للعلوم الاجتماعية الهدف ذاته وطرق البحث ذاتها الخاصة بالعلوم الطبيعية، تدعى "المذهب الطبيعي" أو "التفسير الطبيعي للعلوم الاجتماعية". وقد ينشأ تشويش غالباً، نظراً لوجود نوع من البحث النوعي يدعى "البحث الطبيعي". وفي هذه الحالة فإن مصطلح "طبيعي" يحمل معنى مختلفاً إذ يشير إلى حقيقة أن الدراسات/ البحوث النوعية تجري في وضع "طبيعي" (خلافاً للوضع الذي يتم إنشاؤه بصورة مصطنعة، أي غير طبيعية).

العلمي الذي يشمل اكتشاف القوانين والخضوع إليها – القوانين التي تحكم سلوك العالم المادي، من ناحية، والقوانين التي تحكم السلوك الإنساني، من ناحية أخرى.

يبدأ البحث النوعي من افتراض منهجي مختلف - أي أن موضوع العلوم الاجتماعية أو الإنسانية مختلف أساسا عن موضوع العلوم الفيزيائية والطبيعية وهو بذلك يتطلب هدفاً مختلفا للبحث وجملة طرق مختلفة للاستقصاء.

ويرى الباحثون النوعيون أن السلوك الإنساني محكوم دائما بالإطار الذي يحدث فيه، وان الواقع الاجتماعي (مثلا، الثقافات، والأشياء الثقافية والمؤسسات، وما شابه) لا يمكن تقليصه/ تبسيطه إلى متغيرات بالطريقة نفسها كالواقع المادي، وأن الأهمية الأعظم في النظم الاجتماعية هو فهم وتصوير المضمون الذي يبنيه المشاركون المعنيون في أوضاع أو أحداث اجتماعية معينة. ويسعى البحث النوعي نحو فهم السلوك الإنساني والاجتماعي من منظور "المنتمي للداخل" – أي منظور المشاركين حسبما يعيشون في وضع اجتماعي معين (مثلاً، ثقافة، مدرسة، مجتمع، جماعة، نظام اجتماعي) وهو نوع شخصي جداً من البحث يقر ويعترف "بالفهم الذاتي والتحيزات لكل من المشاركين والباحث في إطار البحث" (Goetz & LeCompte, 1984, p. 95). ويرى المدافعون عن الطرق النوعية، خلافاً لذلك، بأن البحث الكمي معني بشكل رئيسي باكتشاف "الحقائق الاجتماعية" الخالية من الفهم الذاتي للنوايا بمنأى عن الأطر/ السياقات الاجتماعية والتاريخية الخاصة.

تعتمد الأساليب الكمية في العلوم الإنسانية على نموذج افتراضي – استنتاجي للتفسير. ويبدأ البحث بنظرية عن الظواهر المراد بحثها. فمن تلك النظرية يتم استنباط أي عدد من الفرضيات، حيث تخضع بدورها لاختبار باستخدام إجراء مقرر مسبقا كالتصميم التجريبي أو العلي- المقارن أو الارتباطي. أما الهدف الأقصى لاستخدام هذا النموذج الافتراضي- الاستنتاجي فهو مراجعة وإسناد النظريات أو الفقرات الشبيهة بالقوانين الخاصة بالظواهر الاجتماعية والسلوكية القائمة على أساس نتائج اختبار الفرضية.ويتم تشذيب النظريات وتوسيعها (أو استبعادها أحياناً) كي ترتبط بنتائج اختبار مضامينها أو أمثلتها (استنتاجات).

ويعتمد البحث النوعي على نموذج مختلف للتفسير ويعرض أسباباً لدعم هدف مختلف للاستقصاء. فهو يرى، بشكل عام، أن البحث عن العموميات (عبارات شبيهة بالقوانين أو النظريات تظل ثابتة في إطار الزمان والمكان) هو أمر مضلل. فالسلوك الإنساني محكوم دائماً بسياق تاريخي، واجتماعي، وزماني، وثقافي، وعليه فإن نوع القانون وأمثلته للتفسير الذي تتحراه الطريقة الافتراضية – الاستنتاجية مرفوض لصالح نوع الحالات وتفسيراتها (Geertz, 1980). ويسعى الباحثون النوعيون إلى تفسير الأفعال الإنسانية، والمؤسسات الاجتماعية، والأحداث، والعادات، وما شابه، ومن ثم ينشئون تفسيراً أو صورة لما تجري دراسته. إن الهدف الأقصى لهذا النوع من الاستقصاء هو تصوير النمط المعقد لما تجري دراسته بعمق كاف وتفصيل

بحيث يتسنى لمن لم يمارسه أن يفهمه. وعندما يفسر الباحثون النوعيون أو يوضحون مضمون الأحداث والأفعال، وما شابه، فإنهم يستخدمون عموما أحد أنواع التفسير التالية: (1) بناء أنماط من خلال تحليل و إعادة تركيب الأجزاء المكونة، (2) تفسير المعنى/ المضمون الاجتماعي للأحداث، أو (3) تحليل العلاقات بين الأحداث والعوامل الخارجية (McCutcheon, 1981).

تختلف الأساليب النوعية والكمية أيضا في نظرتها إلى دور القيم في البحث. فالباحثون الكميون يعترفون بأن قيم الباحث قد تلعب دورا في تحديد أي الموضوعات أو المشكلات لبحثها، غير أنه يجب أن يكون البحث الفعلي نفسه متحرراً من القيم – أي ينبغي أن يتبع الباحث إجراءات مصممة بشكل خاص لعزل و إبعاد جميع العناصر الذاتية، مثل القيم، عن وضع البحث، بحيث أن ما يبقى هي "الحقائق الموضوعية" فحسب. مثلاً، تصور دراسة تجريبية تشمل صفين مختلفين في السنة الابتدائية الثالثة يكون أحدهما المجموعة التجريبية والآخر المجموعة الضابطة. وتصور أنه جرى وضع الملاحظين/ الراصدين في كل صف لتسجيل التفاعلات بين المدرسين والطلبة. ويفضل الباحثون الكميون ألا يكون الملاحظون مدركين عما إذا كانوا يراقبون المجموعة التجريبية أو الضابطة، وأن يكونوا غير مدركين لخصائص الأفراد (طبقتهم الاجتماعية، معامل الذكاء، والتحصيل الأكاديمي السابق وما إلى ذلك)، وأنهم يستخدمون بروتوكولات/ مراسم ملاحظة منظمة بشكل عال تتطلب استنتاجات متدنية المستوى وتفسيراً قليلاً أو معدوماً، حول ما يحدث في التفاعلات بين المدرس والطلبة. وتستخدم هذه الإجراءات (مثل أصناف تحليل التفاعل لدى فلاندرز المبين في الفصل 7) في بحث كمي لضمان عدم تأثير قيم ومعتقدات الملاحظين على الملاحظات التي يقومون بها. وباتباع هذه الإجراءات للقيام بالرصد، يوفر الباحث الكمي ضماناً قوياً بأن البحث متحرر من القيم.

وخلافاً لذلك، تذهب الطريقة النوعية إلى أن البحث محكوم بالقيم دائماً– ولا يمكن اعتباره متحرراً من القيم على الإطلاق– كما ينبغي أن يكون الباحثون صريحين بشأن الأدوار التي تلعبها القيم في أية دراسة معينة. ويرى الباحثون النوعيون أن البحث محكوم بالقيم في اختيار المشكلة المراد بحثها، وفي اختيار ما إذا كان سيتبنى الطريقة الكمية أو النوعية لمشكلة ما، وفي اختيار الطرق لاستخدامها في بحث تلك المشكلة، وفي اختيار طريقة تفسير النتائج أو الاستنتاجات، وبواسطة القيم المتأصلة في الإطار / السياق الذي تحدث فيه الدراسة (Lincoln & Guba, 1985). ويعتقد الباحثون النوعيون أن من المستحيل تطوير فهم ذي معنى، للخبرة الإنسانية، دون الأخذ بنظر الاعتبار للتفاعل بين قيم وعقائد الباحثين والمشاركين. وعلاوة على ذلك، يرى الباحثون النوعيون أن البحث الإنساني يتطلب تفاعلاً متكرراً ومستمراً وذا معنى بين الباحثين والمستجيبين لهم (الأفراد) وأنه يتعين على البحث أن يضخم ولا يقلص هذا النوع من الصلة (Lincoln & Guba, p. 107). ونظراً لأن البحث النوعي يقر بشكل علني بدور القيم في البحث، ويتطلب تفاعل الباحثين والمستجيبين، فإن ثمة زعماً على نحو متكرر وهو أن الاستنتاجات (النتائج) لمثل هذه الدراسات تظل ببساطة مسالة رأي. ولمواجهة هذا الاتهام،

يستخدم الباحثون النوعيون أساليب مختلفة لعرض مصداقية استنتاجاتهم. (ستناقش هذه الأساليب بتفصيل أكثر فيما بعد).

يلخص (Janesick, 1994, p. 212) خصائص تصميم البحث النوعي:

1- إن التصميم النوعي تصميم شمولي. فهو ينظر إلى الصورة الأكبر، الصورة الشاملة، ويبدأ بالبحث عن فهم للكل.
2- ينظر التصميم النوعي إلى العلاقات ضمن نظام أو ثقافة.
3- يشير التصميم النوعي إلى ما هو شخصي، ووجه لوجه، ومباشر.
4- يركز التصميم النوعي على فهم وضع اجتماعي معين، وليس بالضرورة على إجراء تنبؤات حول ذلك الوضع.
5- يتطلب التصميم النوعي أن يبقى الباحث في سياق وبيئة الأحداث طوال الوقت.
6- يتطلب التصميم النوعي وقتا للتحليل مساويا للوقت في الميدان.
7- يتطلب التصميم النوعي أن يطور الباحث نموذجا لما حدث في السياق والبيئة الاجتماعية.
8- يحتاج التصميم النوعي من الباحث أن يصبح هو أداة البحث. وهذا يعني أنه ينبغي على الباحث أن يمتلك القدرة على ملاحظة السلوك كما عليه أن يشحذ المهارات الضرورية للملاحظة والمقابلة وجها لوجه.
9- يضم التصميم النوعي قرارات الموافقة المعلنة كما أنه حسّاس للاعتبارات الأخلاقية.
10- يضم التصميم النوعي مجالاً لوصف دور الباحث ووصفا لتحيزات الباحث الذاتية وتفضيله الأيديولوجي.
11- يحتاج التصميم النوعي إلى تحليلات مستمرة للبيانات.

الجوانب الحرجة للبحث النوعي Critical Aspects of Qualitative Inquiry

على الرغم أن الباحثين النوعيين يعملون بطرق مختلفة عديدة، إلا انهم يشتركون باهتمام في مجموعة إجراءات خاصة للقيام ببحث معين. وبعض أهم الجوانب الحرجة لطريقة الإجراء تلك، هي ما يلي:

الاهتمام بالسياق: يفترض البحث النوعي أن السلوك الإنساني مقيد/ محكوم بالسياق(*) وأن الخبرة الإنسانية تستمد معناها / مضمونها منه، وعليه فهي لا تنفصل عن التأثيرات الاجتماعية والتاريخية والسياسية والثقافية. وهكذا يتقيد البحث دوماً بسياق خاص أو بيئة ما.

(*) السياق (Context): كلمة تشير إلى جملة الظروف التي تقع ضمنها الأحداث - (المراجع).

ويرى مؤيدو البحث النوعي أن الطريقة الكمية لدراسة التجربة الإنسانية تسعى إلى عزل السلوك الإنساني عن سياقه، فهي تنشغل في تعرية / نزع السياق (Mischler, 1979).

الوضع/ الإطار الطبيعي: ينبغي للسياق الذي تدرس فيه التجربة الإنسانية أن يحدث بشكل طبيعي (كالصف، أو المدرسة برمتها، أو منظمة) وليس بشكل مبتدع أو مصطنع (كالتجربة المختبرية). وهكذا، فإن البحث النوعي يحدث في الميدان ضمن إطار سياق وبيئة كما نجدها. إضافة إلى ذلك، لا يضع البحث النوعي أية قيود مسبقة على ما يدرس. فهو، مثلاً، لا يحدد ولا يعرف ولا يستقصي ولا يختبر العلاقة بين جملة معينة من المتغيرات المستقلة والتابعة بل إنه يدرس الخبرة الإنسانية بشكل شمولي، آخذا بعين الاعتبار جميع العوامل والتأثيرات في وضع معين.

الأداة البشرية: في الدراسات النوعية، يكون الباحث بذاته أداة جمع البيانات. فهو يتحدث مع الناس في إطار السياق والبيئة، ويرصد نشاطاتهم، ويقرأ وثائقهم وسجلاتهم المدوّنة، ويسجل هذه المعلومات في ملاحظات ميدانية وسجل خاص لليوميات. ويعتمد البحث النوعي على طرق العمل الميداني – المقابلة، والملاحظة غير المخططة، وتحليل الوثائق – باعتبارها الوسيلة الرئيسة لجمع البيانات، وتجنب استخدام اختبارات القلم والورقة، والأدوات الميكانيكية وبروتوكولات الملاحظة المنظمة بشكل عال. ويتعامل الباحث النوعي مع البيانات بشكل كلمات وليس بأعداد أو إحصاء، غير أنه من وقت لأخر، قد يجمع بيانات بصيغة عددية. إن إدارة هذا الحجم الكبير من البيانات المولدة من المقابلات والملاحظات وجمع الوثائق يُعد شأنا مهما في الدراسات النوعية.

يحتفظ الباحثون النوعيون بسجل شخصي يسجلون فيه انعكاس أفكارهم، ومشاعرهم، وافتراضاتهم، ودوافعهم، والأسس المنطقية لقراراتهم. وهذه إحدى الطرق التي ينهمك الباحث النوعي من خلالها في قضية البحث المحكومة بالقيم.

التصميم الطارئ/ المتدرج: في الدراسات الكمية، يصمم الباحثون بدقة جميع جوانب الدراسة "قبل" أن يجمعوا أية بيانات فعلية، فهم يحددون المتغيرات ومقاييسها، والإحصاء المزمع استخدامه لتحليل البيانات وما إلى ذلك. وهذا ممكن لان هؤلاء الباحثين يعرفون سلفا ما يبحثون عنه. فلديهم فرضيات محددة أو أسئلة معينة في أذهانهم ويمكنهم أن يتصوروا ما قد يكون عليه اختبار فرضية أو جواب على السؤال. وبغض النظر عن المشكلة أو الظاهرة المعينة التي هي قيد البحث، يصر الباحثون الكميون على أن هذا النوع من تحديد عناصر تصميم الدراسة مهم جدا. وخلافاً لذلك، نادراً ما يحدد الباحثون النوعيون جميع جوانب التصميم قبل بدء الدراسة، بل إن التصميم "ينبثق" مع تكشف ووضوح الدراسة. وهم يكيفون طرقهم وطريقة الإجراء (التصميم) للموضوع القائم. وهذا ضروري إذ إن الباحث النوعي لا يكون أبدا متأكداً عما سيعرفه في وضع معين (لم يقرر الباحث سلفا ما يبحث عنه) لأن ما يمكن

معرفته في وضع معين يعتمد على طبيعة وأنماط التفاعلات بين الباحث والناس والوضع، وتلك التفاعلات لا يمكن التنبؤ بها، ولأن المظاهر المهمة التي تحتاج للبحث لا يمكن معرفتها دوماً إلى أن يشهدها الباحث فعلاً.

وهكذا، فالبحث النوعي يمكن أن يوصف سلفا بطريقة عامة جدا توضح الكيفية التي تتكشف بها الدراسة: فهي "تبدأ بمشكلة بحثية معينة أو سؤال أو موضوع – مثلاً، كيف يواجه العاملون في الحقل الاجتماعي الإجهاد الناشئ عن أعمالهم؟ وكيف يستجيب المدرسون في المدارس الريفية إلى تكنولوجيات التعلم عن بعد؟ ما هو الألم المزمن؟ أو ما هي العلاقة بين المدارس الريفية ومجتمعاتها؟ ويسعى الباحث عندئذ للحصول على سبيل للوصول إلى موقع معين أو مجموعة من الناس يمكن أن يدرس فيها الموضوع ويفاوض في الدخول عن طريق تبني دور كملاحظ – مشارك بالكامل أو مجرد ملاحظ، أو دمج ما بين الاثنين. أن التفاوض على الدخول، وتبني الدور، وبناء وإدامة الثقة مع المشاركين في دراسة ما هي موضوعات معقدة كُتب عنها الكثير (Emerson, 1988). وبعد فترة أولية يألف فيها الموقع، ويحيط علما بالأفراد وما إلى ذلك، يبدأ الباحث بالتركيز على البحث في الجوانب أو القضايا البارزة عن طريق صياغة فرضيات (عملية) أو أسئلة. وهذه الفرضيات أو الأسئلة الأولية يتم تصفيتها (تضييقها) عندما يصبح الباحث اكثر تركيزاً. وخلال هذه المرحلة سيجري الباحث مراجعة لأدبيات الموضوع وذلك من أجل تعميق فهمه للظواهر المدروسة. ويتم في النهاية تثبيت الفرضيات واختبارها حسب تنوع من الإجراءات. وتكون المرحلة الأخيرة في مغادرة الموقع.

المعاينة: تعتبر المعاينة مهمة في البحث النوعي كما الحال في البحث الكمي. فلا يسع الباحثون النوعيون رصد كل شيء قد يكون ذا صلة بمشكلة البحث غير أنهم يحاولون الحصول على عينة الملاحظات التي يعتقد أنها تمثل كل شئ يمكن ملاحظته. وبعبارة أخرى، يختار الباحثون النوعيون عينات هادفة تكفي لتأمين أقصى بصيرة وفهم لما يدرسونه.

ينبغي أن يقرر الباحث أولاً ما هو الشيء الجوهري لدراسة المشكلة. افترض أن دراسة نوعية تركز على الانضباط في نظام مدرسي. في هذه الحالة، يقرر الباحث بشأن أفراد الملاك الذين يريد مقابلتهم (المدراء، المعاونون، المدرسون، الطلبة) وبشأن الأوضاع (صف، ملعب، كافتيريا) لملاحظتها.

يجب أن يطور الباحث خطة لعينة أفراد الملاك والأوضاع التي ستقدم حسب رأيه صورة دقيقة للاتجاهات والأساليب الانضباطية المستخدمة في النظام المدرسي. لقد كتب (Cuba & Lincoln, 1981, p. 276) "أن المعاينة لا تكون، تقريباً، ممثلة أبدا أو عشوائية، إنما هادفة يقصد من دراستها استثمار أراء متنافسة وعلاقات صحيحة جديدة قدر الإمكان. وتتوقف المعاينة عندما تصبح المعلومات فائضة بدلا من أن تتم معاينة الأفراد بصورة ممثلة".

إقامة الثقة: يستخدم الباحثون النوعيون أنواعا من الإجراءات لتحقيق مصداقية البيانات المجمعة وتأكيد تطور فهمهم أو فرضياتهم. ومن بين هذه الأساليب هو المواظبة الطويلة في الموقع والرصد المتواصل لتوفير مدى كاف وعمق للملاحظات. ويعتبر "المسح المثلثي – استخدام مصادر بيانات متعددة، وملاحظين متعددين و/ أو طرق متعددة – أسلوبا آخر يستخدم لتعزيز الاحتمال بأن الفرضيات والتفسيرات صادقة. ففي المسح المثلثي"(*) يستقصي الباحث ما إذا كانت البيانات المجمعة من خلال إجراء أو أداة واحدة تؤكد البيانات المجمعة باستخدام إجراء أو أداة مختلفة. فالمرء يود إيجاد دعم للملاحظات والاستنتاجات بأكثر من طريقة واحدة. فالإيجاز الدوري مع أقران الباحث وتدقيقات الأعضاء (تقديم تفسيرات الباحث لأعضاء في وضع معين لغرض التحقق من صدقها) تعد إجراءات أخرى مهمة.

ولتعزيز موثوقية الدراسة (وهو ما يساوي الثبات تقريباً) فإن الباحث النوعي غالبا ما يجري تدقيقا لأثر المواد التي توثق كيفية إجراء الدراسة بما في ذلك ما تم القيام به، ومتى، ولماذا. ويحتوي تدقيق الأثر على بيانات خام جمعت من مقابلات، وملاحظات، وسجل لقرارات الباحث حول من تجري معهم المقابلة أو ما يجب رصده، ولماذا، وملفات توثيق كيف تم تطوير فرضيات عملية من البيانات الخام التي تم تدقيقها واختبارها واستنتاجات الدراسة، وما إلى ذلك. وباستخدام تدقيق الأثر كدليل، يفحص مدقق مستقل وهو طرف ثالث، دراسة الباحث بغية المصادقة على موثوقية الإجراءات المستخدمة واختبار ما إذا كانت الاستنتاجات قابلة للتأكيد – أي ما إذا كانت مشتقة منطقياً ومستمدة من البيانات المجمعة (Schwandt & Hal, 1988).

التحليل الاستقرائي: في أغلب الدراسات النوعية، يتم جمع البيانات وتحليلها بشكل متزامن. وبعبارة أخرى، لا ينتظر الباحث حتى يتم "الحصول على" جميع البيانات قبل الشروع بتفسيرها. فمنذ بدء المقابلة أو الرصد الأول يتأمل الباحث النوعي في مضمون ما سمعه وما شاهده مطوراً أحاسيسه الداخلية (فرضيات عاملة) حول ما يعنيه، ويحاول تأكيد أو عدم تأكيد هذه الأحاسيس في مقابلات تالية. وتعتبر عملية تحليل البيانات هذه استقرائية – فهي تمضي من البيانات إلى الفرضيات إلى النظرية. وعندما يقلص الباحث البيانات ويعيد بناءها من خلال عملية الترميز والتصنيف، فإنه يتوجه نحو تطوير "نظرية راسية"(**) (grounded theory)، وهي نظرية عن الظواهر المرصودة المرتبطة (الراسية) بشكل مباشر في البيانات الخاصة بتلك الظواهر (Strauss, 1987).

(*) المسح المثلثي (triangulation) يشير إلى أسلوب بحثي لزيادة صدق النتائج باستخدام طرق مختلفة (ومستقلة) في جميع البيانات الخاصة بالمشكلة، وقد يكون ذلك باستخدام أكثر من باحث أو حتى بجمع أنواع مختلفة من البيانات الخاصة بالمشكلة. ولعل جذور التسمية، رياضية، حيث يستخدم هذا المصطلح في طريقة المسح بتقسيم المنطقة إلى مثلثات واستخدام حساب علم المثلثات في حساب عناصر المثلث لحساب المساحة الكلية – (المراجع).

(**) النظرية الراسية (grounded theory): نظرية تخص ظاهرة ما وتنطلق من البيانات المرتبطة بتلك الظاهرة، فهي حسب اجتهادنا "راسية – Grounded" في البيانات حيث يتم إطلاقها (إشهارها) منها – (المراجع).

التقرير: تتنوع، بالطبع، تقارير البحث النوعي حسب طبيعة المنشور الذي تظهر فيه (مثلاً، سيختلف بحث أعد لمجلة البحوث التربوية الأمريكية عن مقالة أو بحث مطول)، ومع ذلك فإنها روائية إلى حد بعيد في صيغتها وتحتوي على أوصاف كثيفة للوضع والسياق كما يتضح في هذا المقطع من دراسة (Alan Peshkin, 1986) الميدانية عن المدرسة المسيحية الأصولية:

> **إن قاعات أكاديمية بيثاني المعمودية النظيفة جدا هي محط اعتزاز وفخر المدير McGraw. فالصفوف المهواة الواسعة في المدرسة الابتدائية حيث تبدو المثيرات غير منظمة في صفوف المدارس الابتدائية في أي مكان. وبخلاف ذلك، فإن صفوف المدرسة الثانوية مقبضة نسبياً رغم تزيينها بالقليل من نشرات المجلات التي تعرض صوراً وأفكاراً جيدة وهي متطلبات الإدارة. أما قاعة الألعاب ذات المنصة التي بنيت على أحد الجوانب فهي واسعة بحجم قاعة الاجتماعات. ومقابل المنصة هناك مطبخ صغير. ولأجل شراء الغذاء يصطف الأطفال أمام جدار قاعة الألعاب، ويلتقطون أواني الطعام، ثم يدخلون إلى قاعة الطعام. وتقع مكتبة المدرسة والغرف الإدارية في الجناح الابتدائي. ومكتب المدير ماكرو المزين بشكل جيد يقع خلف مكتب سكرتيرته التي تعمل كمستخدم متعدد المهام لكل المدرسة – ممرضة، قائدة الاستعراض، مستشارة، مصاحبة عازف بيانو، ومعززة لقواعد المدرسة. وهناك حقيبة مزينة تحتوي على عدة كتب عن ابراهام لنكولن وضعت في مكان إلى جانب حافظة كتب لنكولن نفسه تستند إلى جدار المكتب، وفوقها صورتان مؤطرتان بعنوان غيتسبرغ ولنكولن. وهناك حاجز بارتفاع الخصر يفصل منطقة جلوس الزوار عن المكاتب. وعلى الجدار خلف مقعد الزوار، علقت صورة لثيودور روزفلت وكتاب طبع تحتهه "أفكار ث.ر." وحول هذه الصورة أطباق معدنية: جائزة وليام مولر لطلبة الثانوية، جائزة جماعية للمدرسة الوطنية. وغيرها (ص 33).**

بدلاً من عرض جداول الإحصاء والعروض البيانية للبيانات العددية، تعرض تقارير الدراسات النوعية اللغة الطبيعية للمشاركين في دراسة كما استمدت من بيانات ووثائق المقابلة. مثلاً، في الدراسة ذاتها أعلاه يقدم بيشكن للقارئ التقرير التالي الذي تصف فيه أمينة مكتبة بيثاني ممارستها لعملية الرقابة:

> **... بعض الكتب العلمية، منها ما تطور كثيراً أو انحدر كثيراً في بعض المجالات، غير أن لها صفات جيدة بذاتها فإني أخذها إلى السيد كروكر (مدرس العلوم) وادعه يلتقط ما يمكن أن يستخدمه... أني أبحث عن التطور. وذلك أحد الأشياء. وإني ابحث عن ألفاظ الشتيمة كي استبعدها. لقد وجدت صفحتين عن القرود المتطورة إلى إنسان، وبالطبع نحن لا نوافق مطلقاً على ذلك، لذا قمت بحجبها بالختم ولم انزعج من القراءة على الجانب الأخر لكل منهما. ثم، في البداية، كان هناك فصل عن التطور. فوضعت ذلك بين أقواس بالأحرف السوداء وكتبت (تطور) عليه بحيث أن كل من يقرأ ذلك يعرف**

أنه تطور بدلاً من إتلاف الكتاب برمته، لأن الكثير منه كان جيدا. وإذا وجدت شخصاً عارياً، أرسم سروال سباحة قصير أو أرسم رداءاً صغيراً عليه، لكن ذلك فقط في الكتاب العادي الذي لا يمت بصلة إلى الفن. لكن في الفن، فإن الفن فن، فإذا وجدت شخصا دون ملابس، فهذا ما رسموه. وكان لدينا كتاب قصة واحد حيث جميع الأطفال كانوا يستحمون عراة. ولم يكن ذلك بشيء لذا وضعت سراويل سباحة عليهم.

لقد أصدرنا عشرين كتاباً جديداً حول قيمة الأمانة ومثيل لها. وأعطيت لكل مدرس في الصفوف الأولى، أربعة منها ليقرأها كي يرى ما إذا كانت تحتوي على أي شئ قد يثير قلقنا. أحد هذه الكتب قد استخف بالانضباط، لذا بدلا من العبوس الذي يبدأ على التلميذ الذي عوقب، ولم يقبل بذلك، وضعنا إشارة لاصقة هناك مع وجه مبتسم.

أنواع البحث النوعي

TYPE OF QUALITATIVE RESEARCH

تنطوي أغلب البحوث النوعية على ملاحظة من نوع ما، لكن المدى الذي يشارك فيه الملاحظ في النشاط الذي تجري ملاحظته، فإنه يختلف. فالبحث النوعي يمكن تصنيفه إلى صنفين رئيسيين: ملاحظة المشارك، وملاحظة غير المشارك. أما البيانات الحاصلة من هذه الإجراءات فتتكون أساساً من أوصاف لفظية للتفاعلات الاجتماعية والسلوك بدلاً من أرقام وإحصائيات هي نموذجية في البحث الكمي.

رصد المشارك Participant Observation

في "رصد / ملاحظة المشارك" يدرس الباحث مجموعة معينة بحيث يصبح جزءاً من المجموعة. يلاحظ ويقابل، ويشارك فعلا في نشاطاقم. إن دور الباحث كملاحظ قد يكون أو لا يكون معروفا لدى الناس المرصودين. وفي بعض الحالات، يصبح الراصد/ الملاحظ عضواً مشاركاً بالكامل في المجموعة ويحاول مشاطرتها في خبراتها، ويختفي دوره كباحث عن المجموعة. فمثلاً قد ينتحل الباحث دوراً للشخص المشرد في مدينة كبيرة لكي يعرف عن مشاعر، وعلاقات، ومشكلات المجموعة.

في حالات أخرى، ينضم الباحثون جهاراً بمجموعة معينة لهدف معلن، هو دراسة المجموعة. ويشترك الباحث في المجموعة حيث يكون وضعه كملاحظ/ باحث معروف لدى الخاضعين للدراسة. مثلا، يستطيع الباحث التركيز على دراسة صف للتدريب المهني لأفراد الرعاية الاجتماعية، أو جمعية معينة مثل جمعية المدمنين على الكحول المجهولين.

سواء كان هدف الباحث معروفاً لدى المجموعة أو مكتوماً فإن ذلك يعتمد على الموقف.

فباحث راشد يدرس عصابة مراهقين، قد لا يمكنه إخفاء هدفه، بينما يمكن للمرء أن يدعي بأنه شخص مشرد. ومن اليسير طرح أسئلة وتسجيل ملاحظات إن كان أفراد المجموعة يعرفون هدفك، وعلاوة على ذلك، قد يكون من المناسب أخلاقياً أكثر جعل الناس يدركون ما يجري. لكن الصراحة، من ناحية أخرى، قد تولد مشكلات. فالمجموعة التي تعرف أنها مرصودة، قد يتصرف أفرادها بشكل مختلف عما هم عليه عادة، أو قد لا يكونون صادقين عند الإجابة على الأسئلة. ويمكن أن يؤدي ذلك إلى صورة غير دقيقة للمجموعة وتفاعلاتها.

إن للملاحظة المشارك مزيّة توفير صورة مفصلة وشاملة، غير أنها عرضة لمشكلات تورط الراصد عاطفياً في المجموعة مما يفقده الموضوعية.

ملاحظة/ رصد غير المشارك Nonparticipant Observation

في رصد غير المشارك يرصد الباحث دون أن يشارك في النشاط المرصود.وهناك ثلاثة أنواع شائعة الاستخدام في الرصد لغير المشارك وهي: الرصد الطبيعي ودراسات الحالة وتحليل المضمون / المحتوى.

الرصد الطبيعي Naturalistic Observation

في هذا النوع من البحث النوعي، يقوم الباحث بمجرد رصد وتسجيل الأحداث كما وقعت بشكل طبيعي. ولا تتم أية محاولة لتغيير الوضع بأي حال، ولأن الذين يتم رصدهم غير مدركين للرصد، فإن سلوكهم لا يتغير بسبب وجود الباحث.

فالباحث الذي يرغب في دراسة السلوك العدواني لدى أطفال ما قبل سن المدرسة، قد يختار الرصد الطبيعي كطريقة بحث. وبوسع الباحث استخدام مرآة ذات اتجاه واحد للرؤية، أو كاميرا خفية، أو أي أسلوب آخر غير ظاهر لرصد سلوك الأطفال. ويستطيع الباحث أن يرصد الأطفال في ساحة اللعب، عن بعد، بحيث لا يكون حضوره ملحوظاً لدى الأطفال.

قد يستغرق الرصد الطبيعي البسيط وقتاً طويلاً إذ على المرء أن ينتظر حدوث السلوك بشكل طبيعي. ولهذا السبب، قد يصطنع بعض الباحثين موقفاً طبيعياً معداً لاستخلاص السلوك المراد رصده. ورغم أنه مصطنع، فإن الباحث يسعى إلى الحفاظ على واقعية الموقف ويقوم ببعض محاولات الرصد بطريقة غير ملحوظة لدى الأفراد. فقد استخدم (Hartshorne & May, 1928) رصداً طبيعياً معداً في دراستهما الكلاسيكية للغش في الصف (انظر الفصل 7).

دراسات الحالة Case Studies

إن كلاً من دراسات الحالة وتجارب الفرد الواحد (انظر الفصل 9) تدرس كل منها فرداً واحداً أو وحدة اجتماعية واحدة مثل العائلة، أو النادي، أو العصابة. وعلى أية حال، تركز

تجارب الفرد الواحد على سلوك واحد أو عدد محدود جداً من أنواع السلوك بينما تحاول دراسات الحالة وصف المدى لأنواع سلوك الفرد وعلاقتها بتاريخ وبيئة الفرد. وفي تجربة الفرد الواحد، يدخل الباحث معالجة محددة بهدف دراسة تأثير هذه المعالجة على الفرد، بيد أن الباحث في دراسة الحالة يرصد رد فعل الفرد إزاء أحداث تقع بصورة طبيعية.

وفي دراسة الحالة، يسعى الباحث إلى تفحص فرد أو وحدة بعمق. فالباحث يسعى إلى اكتشاف جميع المتغيرات المهمة في تاريخ أو تطور الفرد. ويكون التأكيد على فهم سبب قيام الفرد بما يقوم به وكيف يتغير سلوكه عندما يستجيب للبيئة. ويتطلب هنا دراسة مفصلة لفترة طويلة من الزمن. ويجمع الباحث البيانات حول الحالة الراهنة للفرد، وخبراته السابقة، والبيئة، وكيف تتصل هذه العوامل مع بعضها.

تنشأ العديد من دراسات الحالة من محاولات حل المشكلات. لقد بدأت دراسات الحالة المشهورة لفرويد في سعيه لمساعدة مرضاه في حل مشكلاتهم الشخصية. فعندما حاول الغوص بعمق في ديناميكية(*) شخصيات مرضاه رأى أن العلاقات التي رصدها بينهم وبين بيئاتهم قد تكون مميزة لأشخاص آخرين لهم مشكلات مشابهة. وقد نشر تقارير مفصلة عن مقابلاته مع مرضاه وتفسيراته لأفكارهم، وأحلامهم، وأفعالهم، على افتراض أنه يمكن لهذه الدراسات أن تؤدي إلى تعميمات بعيدة الأثر.

إن أعظم فائدة لدراسة الحالة هي احتمال العمق، فهي تسعى إلى فهم الطفل بكليته أو الراشد بكليته في الحاصل الكلي لبيئة الفرد. وهذا الفهم المتعمق لا يتصل بأفكار الفرد الراهنة فحسب، بل بماضيه وبيئته وعواطفه وأفكاره. ويحاول الباحث تحديد "سبب" تصرف الفرد كما يفعل، وليس مجرد تسجيل سلوكه. وتوفر دراسات الحالة في الغالب الفرصة للباحث لتطوير فهمه للجوانب الأساسية للسلوك الإنساني. إن التفحص المتعمق المكثف الذي يميز هذا الأسلوب قد يؤدي إلى اكتشاف علاقات لم تكن موضع شك سابقا.

من ناحية أخرى، تعتبر مزايا دراسة الحالة بأنها أيضاً نقاط ضعفها. فرغم أنها تتمتع بالعمق، إلا أنها تفتقد حتما إلى العرض/ الاتساع. كما أن القوى المحركة (الديناميكية) لفرد أو وحدة اجتماعية قد يكون لها علاقة قليلة بديناميكية الآخرين. وفي الواقع فإن أغلب دراسات الحالة تنشأ عن الإرشاد، أو الجهود العلاجية، وهي بذلك توفر المعلومات عن الأفراد الاستثنائيين وليس الممثلين.

إن فرص الفهم العميق في دراسة الحالة هي كذلك فرص للذاتية أو حتى المحاباة. ويمكن أن تقرر المفاهيم المسبقة للباحث أنواع السلوك التي تجري ملاحظتها وأنواع السلوك التي يتم تجاهلها وكذلك الطريقة التي تفسر بها الملاحظات.

(*) الديناميكية (dynamics)؛ مصطلح يشير إلى وجود قوى محركة خلف الظاهرة – (المراجع).

لقد عانت سمعة طريقة دراسة الحالة لأن بعض الباحثين في الماضي فسروا ملاحظاقم في مفاهيم بنائية يستحيل تأكيدها أو رفضها من خلال الدراسة التجريبية.

وحيث أن المدى، الذي تستطيع فيه دراسات الحالة تقديم تعميمات صادقة، محدود فإن فائدتها الكبرى ليس باعتبارها أدوات لاختبار الفرضيات بل في إنتاج الفرضيات التي يمكن بعدئذ اختبارها من خلال استقصاء أكثر دقة. مثلاً، الفهم الذي حصل عليه جان بياجيه، في دراساته للحالة الشهيرة حول نضوج الفكر، قدم فرضيات مفيدة جرى بحثها منذ ذلك الحين خلال طرق أخرى.

في الأمثلة التي تنشأ فيها الدراسات الميدانية من محاولات معرفة الأشخاص بغية مساعدتهم، فإن الجوانب البحثية للدراسة تأخذ المكانة الثانية. ومع ذلك فإن دراسات الحالة تجري غالباً كذلك مع هدف أولي للحصول على المعرفة. فدراسة (Itard) الكلاسيكية للحالة على "الولد الشقي من أفيرون" (1962) كانت جهدا قيما لمعرفة تأثيرات الحضارة من خلال دراسة ولد ترعرع بمعزل عن الحضارة في فرنسا في القرن الثامن عشر. ولقد أجريت دراسات بياجيه للحالة بهدف معرفة النمو العقلي لدى الأطفال بدلا من إفادة الأفراد الذين تضمنتهم الدراسة.

تحليل المحتوى/ المضمون Content Analysis

يعتبر تحليل المحتوى طريقة بحثية مطبقة على مواد مكتوبة أو مرئية بهدف التعرف إلى خواص محددة للمادة. ويمكن للمواد التي تجري تحليلها أن تكون كتباً مدرسية، أو صحفاً، أو خطابات، أو برامج تلفزيونية، أو إعلانات، أو مقطوعات موسيقية، أو أية مجموعة من الوثائق. ويشيع استخدام تحليل المحتوى في التربية. فمثلاً، قد يقوم باحث بتحليل كتب التاريخ في مدرسة ثانوية في مقاطعة دراسية معينة ليرى مدى تكرار ذكر النساء ومقدار النقاش المعطى لكل منهن. ويمكن للمرء أن ينظر إلى العمل الكتابي للتلاميذ لتصنيف أخطاء الهجاء أو النحو وطبيعتها وترددها / تكرارها.

وقد يجري تحليل المحتوى في إطار تصميم طارئ / تدريجي. أو قد تجري مثل هذه التحليلات في إطار البحث الكمي مع متغيرات محددة مسبقاً وأعداد تتولد لتمكين الباحث من الوصول إلى استنتاجات حول هذه المتغيرات المحددة. مثلاً بحث (Allen, Allen & Sigler, 1993) حول تنميط دور الجنس (ذكر، أنثى) في شريحة لأدبيات الأطفال وهي على وجه التحديد، الكتب التي فازت بجائزة كالديكوت ميدال. وكان هدف التحليل هو تحديد ما إذا كانت هذه الكتب قد قدمت نمطية لأنواع السلوك حسب الجنس، للأطفال من خلال الشخصيات في النص والصور. لقد قاموا بمقارنة الكتب من عام 1938-1940 ومن 1986-1988 في أحد عشر صنفاً حيث يمكن أن يحدث تنميط حسب الجنس - تشمل الشخصيات في النص والصورة، ومهنة الشخصيات الرئيسية، وما إذا كانت الشخصيات فعالة أو خاملة، منطوية أو متفتحة، ذات

أدوار تقليدية أو غير تقليدية، وما إلى ذلك. وقد وجدوا اتجاهاً ضعيفاً نحو التمثيل المتساوي في سبعة من الأصناف الأحد عشر على الرغم أن الذكور كانوا لا يزالون هم الغالبية في الشخصيات في كل صنف. لقد أتصف الذكور بالفاعلية، والانفتاح، وغير التقليدية، والمهن المتنوّعة أكثر من الإناث في كلا الفترتين الزمنيتين. لقد استنتج الباحثون أن التنميط حسب الجنس قد تقلص، لكنه بقي سائداً في كل صنف من البحث.

توضح الدراسة أعلاه الخطوات المشمولة بتحليل المحتوى:

1- تحديد الظاهرة المراد بحثها (مثل التنميط "Stereotyping" حسب الجنس).

2- اختيار الوسط الإعلامي الذي ستؤخذ الملاحظات منه (مثل كتب جائزة كالديكون ميدال لفترات زمنية محددة).

3- صياغة أصناف ترميز شاملة ومنفصلة تبادلياً بحيث يمكن إحصاء المحتوى اللفظي أو الرمزي (كالأصناف الأحد عشر حيث يمكن للتصنيف حسب الجنس أن يحدث).

4- التقرير بشأن خطة المعاينة المستخدمة بغية الحصول على عينة ممثلة للوثائق (مثل العينة المكونة من جميع الفائزين بجائزة كالديكون لفترات من سنتين). وقد يقرر أحدهم النظر إلى ثلاثة إصدارات في الأسبوع لصحيفة على مدى سنة واحدة، مثلاً، أو كل إصدار من مجلة أسبوعية لسنة.

5- تدريب المشفرين / المرمّزين بحيث يمكنهم تطبيق نظام التشفير / الترميز الذي تم تطويره مما يساهم في ثبات تحليل المحتوى. وبعض التشفير لا يكون مباشراً، بل قد يحتاج إلى استنتاجات حول ما إذا كانت مجموعة من الأقليات، مثلاً، يجرى تصويرها بشكل إيجابي أو سلبي. وينبغي أن يكون بعض المشفرين قادرين على تشفير / ترميز الوثائق باستخدام البرنامج للحصول على نتائج متّسقة. وإذا كانت تقديرات الثبات مرضية، يمكن للمرء عندئذ المضي إلى الخطوة التالية. وإذا كانت أقل من مرضية، فقد يكون المشفرون مفيدين في مراجعة تعريفات التصنيفات لجعلها أوضح وأكمل.

6- تحليل البيانات، التي قد تنطوي على مجرد التكرارات والنسب المئوية في الأصناف المختلفة. وقد تكون هذه الخطوة بطيئة وتستغرق وقتاً. ولحسن الحظ يمكن للحاسوب الآن أن ينفذ تحليل المحتوى بسرعة ودقة. فلدى العديد من الجامعات أجهزة مسح / تدقيق صوري (image scanner) مثل (Kurzwell Data Entry Machine) التي يمكنها أن تقرأ الصفحات المطبوعة، وتحول النص إلى قرص. ويمكن استخدام برنامج الحاسوب للبحث في النص المدروس بدقة / الممسوح (scanned) وإيجاد كلمات وعبارات تنطبق على المعايير المحددة. وبوسع المرء أن يحصل على قائمة بالكلمات المحددة، مثلاً، والتكرار الذي تظهر فيه الكلمات في الوثائق المطبوعة. حتى أن برامج حاسوبية أكثر تطوراً، وتشمل على تلك

القادرة على تصنيف المحتوى وتفسير معاني الكلمات حسب النص، يمكن توقع توفرها حسبما تمضي البحوث نحو المجال المعروف "بالذكاء الاصطناعي".

وتكمن مزية تحليل المحتوى في كونه لا يجلب الانتباه. فحضور الراصد لا يؤثر على ما يرصد. ولا يحتاج المرء إلى إدراج تعاون الأفراد أو الحصول على إذن بإجراء الدراسة. ومزية أخرى لتحليلات المحتوى هي سهولة تكرارها.

أنواع أخرى للبحث النوعي
Other Type of Qualitative Research

ثمة أساليب أخرى بالوسع استخدامها كجزء من دراسات ملاحظة المشارك أو دراسات الحالة، أو بالوسع استخدامها بحد ذاتها:

المقابلات المركزة Focused Interviews

المقابلة المركزة طريقة لجمع البيانات النوعية عن طريق طرح أسئلة على الأفراد حول سلوكهم. والمقابلة المركزة أكثر مرونة وانفتاحاً من مقابلة المسح التي نوقشت في الفصل 12. فالمستجيبون أحرار في الإجابة بعباراتهم ويمكنهم الإجابة بشكل مختصر أو بشكل مطول. وقد تختلف الأسئلة المطروحة من شخص لآخر. وتسجل الإجابات عن طريق اخذ الملحوظات خلال المقابلة أو بعدها مباشرة أو بشريط سمعي.

مثلاً، في دراسة نوعية مصممة لفهم مواظبة الطلبة السود في جامعة كان غالبيتها من البيض، قابل (Craft, 1991) ثلاثة و أربعين طالباً أسود حول إدراكهم للخبرة الأكاديمية، وأسس النجاح الأكاديمي، وسبب كون بعض الطلبة السود الذين يلتحقون في جامعة يهيمن عليها البيض هم أكثر نجاحاً من الآخرين. ودلت إجاباتهم على أن القائمة الاعتيادية من العوامل العلّية مثل القدرة، والجهد، وصعوبة الواجب، تعطي صورة جزئية فقط عن الكيفية التي يقيّم بها الطلبة الأداء الأكاديمي. وشدد هؤلاء الطلبة على أهمية معتقدات الطلبة من الأقليات عن الكيفية التي يفهمهم بها أعضاء هيئة التدريس وأقرانهم الطلبة.

استثمر (Kagan, Dennis, Igou & Moore, 1993) المقابلات المركزة بهدف اختبار تأثيرات برنامج تطور الموظفين على الحياة المهنية لأربعة مدرسين من المدارس الابتدائية كانوا قد شاركوا فيه. وكان البرنامج قد اخذ مدرسي الابتدائية المتمرسين إلى كلية التربية في جامعة لفصول دراسية مدتها سنتان عملوا خلالها كمساعدين مؤقتين في الكلية. وقال المدرسون الذين جرت مقابلتهم أنهم لم يتعلموا بشكل جذري أشياء جديدة في البرنامج، لكن ذلك مكنهم من إيضاح ما كانوا يعرفونه عن التدريس وإعادة إلزام أنفسهم بدور مهنة المدرس. لقد أوضح هذا البحث النوعي عن إمكانية الزمالة ما بين المدرسة والجامعة لتعزيز الحياة المهنية للمدرسين.

البحث الأثنوجرافي Ethnographic Research

لقد تم تطوير طريقة البحث الأثنوجرافي من قبل علماء الأنثروبولوجيا(*) (مثل Margaret Mead) كطريقة لدراسة ووصف الثقافات الإنسانية. فعلماء الأنثروبولوجيا يدمجون أنفسهم في حياة الناس الذين يدرسونهم باستخدام ملاحظة المشارك الموسعة والمقابلة المعمقة للحصول على صورة شاملة للمجموعة. ويستخدم مصطلح (أثنوجرافيا) للإشارة إلى دراسة ثقافة ما مع الناتج النهائي للبحث. لقد انتقلت الأثنوجرافيا من الأنثروبولوجيا إلى حقول معرفة أخرى، مشتملة على التربية، حيث ساهمت بمنظور بحثي يؤكد على الطريقة الشمولية في جمع البيانات ودراستها بالإضافة إلى أسلوب استقرائي لتحليل البيانات. وقد قدمت معلومات أساسية حول طبيعة النقل/ التواصل الثقافي والعملية النظامية للتربية التي أدت إلى الفهم المتعمق والتعميمات المفيدة للمارسيين التربويين (Chilcott, 1992).

وتعتبر دراسة Peshkin مثالاً للبحث الأثنوجرافي في التربية. ففي دراسة عام 1991 (لون الغرباء ولون الأصدقاء) عرض بيشكن دراسة معمقة أخذت عاماً عن ثانوية Riverview وهي مدرسة في مجتمع طبقة عاملة في كاليفورنيا بها أعداد كبيرة من الطلبة جاءوا من جماعات عرقية مختلفة. لقد استخدم ملاحظة المشارك والمقابلات لاستكشاف دور العرقية في هذه المدرسة ذات الأعراق المتعددة. ومع أن Riverview مرت باضطرابات عرقية وعنف قبل بضع سنوات، إلا أن بيشكن وجد أن المدرسة كانت تتسم بسلام عرقي. لقد وصف المدرسة بكونها قصة نجاح اجتماعي- وهي مكان حيث لم يتم فصل الغرباء عن الأصدقاء على أساس اللون. إن التفاعلات الاجتماعية التي تحدث عادة ضمن جماعة عرقية فقط، امتدت عبر كل الجماعات العرقية في هذه المدرسة. إن مستوى النجاح الأكاديمي للطلبة، على أية حال، لم يناظر النجاح في التفاعلات الاجتماعية. لقد وجد بيشكن مستوى متدنياً وبشكل متواصل للنجاح الأكاديمي خصوصاً بين الطلبة السود والطلبة المنحدرين من اسبانيا أو البرتغال أو أمريكا اللاتينية (Hispanic students).

في دراسة أثنوجرافية أخرى بعنوان (الاتحاد غير التام: الدمج المدرسي(**) وصراع المجتمع) عام 1982، ركز Peshkin على المقاطعة الدراسية الريفية في الوسط الغربي التي كانت تقاوم الأمر الرسمي للدمج. وبعد دراسة دقيقة للمقاطعة الدراسية وتاريخها، أوضح بيشكن أنه عندما تدرس الأمر من خلال وجهات نظر الناس في المقاطعة تجد كيف كان سلوكهم معقولاً ولم يكن "غير عقلاني" مطلقاً كما أشير. هذا ما يحاول البحث الأثنوجرافي عمله – بناء فهم كامل لمجموعة معينة من منظورات أفراد تلك المجموعة. وهناك بحث أثنوجرافي آخر قدم فهماً متعمقاً عن الأطفال المهاجرين والأقليات العرقية في مدارسنا.

(*) الأنثروبولوجيا (Anthropology): تشير إلى دراسة الجنس البشري، وخاصة أصوله، وتطوره، وعاداته ومعتقداته. وتترجم إلى "علم الإنسان" – (المراجع).

(**) دمج مدرسي (School consolidation)، يشير إلى دمج عدة مدارس صغيرة في مدرسة واحدة تدعى المدرسة المدمجة/ الموحدة – (المراجع).

منهجية / طريقة البحث Methodology

يحدد (Spradley, 1980) سلسلة الخطوات التي تكوّن منهجية / طريقة البحث الأثنوجرافي.

1- **اختيار مشروع أثنوجرافي:** قد يتنوع مدى هذه المشروعات بشكل كبير من دارسة مجتمع معقد بكليته مثل مجموعة الإسكيمو للصيد في الاسكا إلى موقف اجتماعي واحد أو مؤسسة واحدة مثل نادٍ في المدينة أو ملعب في مدرسة. وسيكون الباحث المبتدئ حكيماً في تحديده لمدى مشروعه إلى موقف اجتماعي معين بحيث يمكن إكماله في وقت معقول. إن للموقف الاجتماعي دوماً ثلاثة مكونات: المكان، والعاملون، والنشاطات.

2- **طرح أسئلة أثنوجرافية:** يحتاج الباحث إلى أسئلة في ذهنه ترشده إلى ما يسمعه ويراه وجمع البيانات.

3- **جمع بيانات أثنوجرافية:** يقوم الباحث بالعمل الميداني ليكشف عن نشاطات الناس، والخصائص المادية للموقف، وما يبدو عليه الأمر حين يكون جزءاً من الموقف. وتبدأ هذه الخطوة عموماً بنظرة تضم الملاحظات الوصفية الواسعة. وبعد الإمعان في البيانات ينتقل الباحث إلى ملاحظات أكثر تركيزاً. وهنا يستخدم ملاحظة المشارك، والمقابلات العميقة، وما إلى ذلك لجمع البيانات.

4- **وضع سجل أثنوجرافي:** تشمل هذه الخطوة ملحوظات ميدانية وصور، ووضع خرائط، واستخدام أية وسيلة أخرى لتسجيل الملاحظات.

5- **تحليل البيانات الأثنوجرافية:** يتبع العمل الميداني دوماً تحليل للبيانات يؤدي إلى أسئلة جديدة وفرضيات جديدة، وجمع بيانات وملحوظات ميدانية أكثر وتحليل أكثر. وتستمر هذه الدورة حتى يكتمل المشروع.

6- **كتابة البحث الأثنوجرافي:** ينبغي أن يكتب هذه البحث بحيث يتم فيه إحياء ثقافة أو مجموعة معينة، وجعل القراء يشعرون انهم يفهمون الناس وطريقة عيشهم. وقد يمتد التقرير الأثنوجرافي في الطول من عدة صفحات إلى مجلد أو أثنين. ويمكن للمرء أن يبسط هذه المهمة جداً من خلال البدء في الكتابة مبكراً أثناء ما تتراكم البيانات بدلاً من الانتظار حتى النهاية. وستكون مهمة الكتابة أيسر أيضاً إذا ما قام المرء قبل الكتابة، بقراءة البحوث الأثنوجرافية الأخرى المكتوبة جيداً.

البحث التاريخي HISTORICAL RESEARCH

البحث التاريخي محاولة لتوطيد الحقائق والتوصل إلى استنتاجات تخص الماضي. فالمؤرخ يحدد الأماكن ويقيّم، ويفسر الدليل الذي يمكننا من خلاله معرفة الماضي بشكل منتظم وموضوعي. ووفقاً للدليل القائم، تستمد الاستنتاجات الخاصة بالماضي من أجل زيادة معرفتنا حول كيفية وسبب وقوع الأحداث الماضية والعملية التي أصبح الماضي من خلالها حاضراً. وتكون النتيجة المأمولة فهماً متزايداً للحاضر وأساساً منطقيا أكثر لصنع الخيارات.

يعمل المؤرخ تحت عوائق مختلفة عن تلك التي يواجهها الباحثون في الميادين الأخرى. فالضبط في المعالجة، والقياس والمعاينة محدودة، وليس هناك فرصة للاستنساخ. وكما هو الحال في البحث الوصفي والعلّي – المقارن فإن المتغيرات المستقلة أو متغيرات المعالجة لا تخضع لضبط الباحث. وتنطبق جميع أنواع الحذر في تفسير هذه الدراسات أيضاً على البحث التاريخي. وعلى أية حال، فإنه يمكن في البحث الوصفي والعلّي – المقارن، ضبط القياس عادة من خلال القرار حول المقاييس التي سيجري اعتبارها متغيراً تابعاً. أما المؤرخون، فليس لديهم أي خيار بشأن الوثائق، والآثار، وما صنعه البشر التي بقيت مع مرور الزمن، فلديهم سيطرة محدودة على نوع الأسئلة التي سيطرحونها عن هذه المصادر والمقاييس التي سيطبقونها عليها. فعند مقابلة شهود الأحداث الماضية ولدى البحث عن سجل تاريخي يستطيع الباحثون اتخاذ القرار بشأن الأسئلة التي تطرح وما الذي سيجرى قياسه. لكنهم يستطيعون فقط قياس تلك الأشياء التي يتذكرها الشهود أو ما تحويه السجلات.

في البحث الوصفي والتجريبي، يمكن للباحثين محاولة ضبط المعاينة، أي يمكنهم أن يقرروا لأنفسهم من سيدرسون. فبوسع المؤرخين أن يدرسوا، فقط، الناس الذين بقيت سجلاتهم وما صنعه البشر. فإذا أهملت الصحف شريحة معينة من المجتمع ولم تتوفر هناك أية مصادر أخرى لذلك المجتمع، فعندئذ يعجز المؤرخون عن التقييم المباشر للإسهامات التي قدمتها تلك الشريحة من السكان لحياة ذلك المجتمع. وثمة قيد آخر يصطدم به الباحثون التاريخيون حيث يتعذر عليهم وضع افتراض لمجرد عدم وجود سجل، وعلى العكس لا يمكن الافتراض بأن مؤامرة الصمت قد شوهت السجل التاريخي.

المصادر الأولية والثانوية Primary and Secondary Sources

يصنف المؤرخ المواد على أنها "مصادر أولية وثانوية"، فالمصادر الأولية هي وثائق أصلية أو آثار، أو أدوات بشرية. وهذه هي النتائج المباشرة للأحداث أو سجلات شهود العيان. والأمثلة على ذلك هي محاضر اجتماع المجالس المدرسية، شريط فيديو أصيل للعبة كرة السلة، مجموعة أعمال فنية أكملها طلبة الصف الثالث. في المصادر الأولية يتطفل ذهن الملاحظ فقط بين الحادث الأصلي والباحث. لاحظ أن ذهن المراقب / الملاحظ يأتي بين الحادث والسجل في كل من أمثلتنا.

فشخص ما قد قرر ما سيتعين أو لا يتعين تسجيله في محاضر اجتماعات مجلس المدرسة، ومتى تعمل أو تتوقف الكاميرا وأين تركز في لعبة كرة السلة، و أي الأعمال الفنية يجب الإبقاء عليها.

وفي المصادر الثانوية يأتي أيضاً ذهن غير الملاحظ بين الحادث ومستخدم السجل. فإذا كان مراسل صحيفة حاضرا في اجتماع مجلس المدرسة، فإن تقرير المراسل سيكون مصدراً أولياً. فإذا اعتمد المراسل على محاضر الاجتماع أو مقابلة مشارك لإعداد التقرير، فإن التقرير عندئذ سيكون مصدراً ثانوياً. أما الأمثلة المألوفة عن المصادر الثانوية فهي كتب التاريخ ومقالات الموسوعات، ومراجعات البحوث. ويحاول المؤرخون استخدام المصادر الأولية متى ما أمكن ذلك.

النقد الخارجي والداخلي External and Internal Criticism

لقد برهنت فكرتان على فائدة في تقييم المصادر التاريخية، وهما مفهوما النقد الخارجي (أو الأدنى) والنقد الداخلي (أو الأعلى).

يتساءل النقد الخارجي أساسا عما إذا كان الدليل قيد الدرس هو دليل موثوق به، وبالاعتماد على طبيعة الدراسة، قد ينطوي على أساليب كالمصادقة على صحة التواقيع، أو التحليل الكيميائي للأصباغ أو الكاربون في تحديد تاريخ الأدوات. افترض أن لدى مؤرخ رسالة تصف مدارس مساتشوستس حيث جرى الاعتقاد بأن هوريس مان كتبها. وباستخدام النقد الخارجي يسأل الباحث، هل الورقة من العصر الصحيح؟ هل هي كتابة مان؟ هل تتطابق وجهة النظر وأسلوب الكتابة مع كتابات مان الأخرى؟

وبعد التثبت من موثوقية دليل ما، يمضي المؤرخ الباحث نحو النقد الداخلي الذي يتطلب تقييماً لجدارة الدليل- مثلاً، ما إذا كانت الوثيقة تقدم تقريراً صحيحاً لحدث معين. ويمكن الإجابة على مثل هذا السؤال بمقارنته مع الأسئلة الأخرى التي تلقي الضوء على حدث معين أو توفر معلومات إضافية عن حادث معين والناس أو الظروف المحيطة به. ففي المثال، قد يسأل الباحث، هل أن وصف مان للمدارس غير متحيز؟ هل يتفق هذا الوصف مع الأوصاف المعاصرة الأخرى للمدارس؟

وبما أن هناك قيود على البحث التاريخي، فإن بوسع المرء أن يطرح سؤالاً عن سبب اللجوء إليه. والسبب الأساسي هو عدم وجود طريقة أخرى لاستقصاء العديد من الأسئلة. فكيف يتسنى لأحدهم تقييم نتيجة حوادث إطلاق النار في ولاية كينت وغيرها من الاضطرابات في المباني الجامعية في ربيع عام 1970؟

إن إحدى فوائد البحث التاريخي، وأحيانا سبب استخدام هذه الأسلوب، تكمن في كونه غير ملحوظ. فالباحث ليس مشاركاً مادياً بالموقف المدروس. وليس هناك خطر من تفاعل

القائم بالتجربة – الفرد، كما أنه ليس هناك حاجة للحصول على إذن من سلطات المدرسة لإجراء البحث. فالمؤرخ يحدد مواقع الوثائق المناسبة، ويجمع البيانات الملائمة ويستمد الاستنتاجات عن بعد من الموقف الذي تجري دراسته. إضافة إلى ذلك، قد يقدم البحث التاريخي منظورات جديدة إلى وضع مأزوم. فالطبيعة غير المشاركة للبحث التاريخي قد تجعله مقبولاً في موقف مشحون عاطفياً حيث تكون فيه الأنواع الأخرى للبحوث مستحيلة.

وبسبب قيود البحث، فيجب، على أي حال، توخي الحذر في تعميم نتائج البحث التاريخي. وينبغي على الطلبة الذين يخططون لإجراء دراسة تاريخية العودة إلى المراجع والمصادر المناسبة حول طرق البحث التاريخي.

الخلاصة SUMMARY

يشمل البحث النوعي مجالاً من الإجراءات أهمها الملاحظة، والمقابلات، والتحليل الوثائقي. والدراسات النوعية نمط مميز من البحث في التربية والعلوم الاجتماعية حيث يمكنها أن توفر تقارير حيوية وثرية ومفصلة عن التجربة الإنسانية. وتستند هذه الدراسات على طريقة مختلفة جوهرياً في دراسة الواقع الاجتماعي عن تلك الأسس التي ترتكز عليها الطريقة الكمية الشائعة في دراسة التربية. فالبحوث النوعية تتطلب مجموعة مهارات لا يتم بسهولة تعلمها بدراسة محتويات الكتب عن الطرق. فخبرة العمل الميداني في التفاوض بشان الوصول إلى موقع معين، وتطوير دور الباحث، وإقامة الثقة والحفاظ عليها مع المشاركين في الدراسة، وإجراء وتسجيل المقابلات والملاحظات، وإدارة البيانات، والقيام بتحليل البيانات تعد ضرورية لكي يصبح كفؤا في هذه الطريقة من البحث.

وتستخدم سجلات الأحداث غير الراهنة والآثار في البحث التاريخي لتوليد فرضيات واختبارها. وتستخدم المصادر الأولية بأكبر قدر ممكن. ويسعى الباحث التاريخي إلى تأكيد موثوقية المصادر من خلال النقد الخارجي وصدقها من خلال النقد الداخلي. وبسبب نقاط الضعف المتأصلة، ينبغي توخي الحذر البالغ في تعميم النتائج التي تم التوصل إليها من خلال البحث التاريخي.

مفاهيم أساسية Key Concepts

تدقيق الأثر	audit trail
دراسة الحالة	case study
اهتمام بالسياق	concern for context
تحليل وثائقي	documentary analysis

emergent design	تصميم طارئ / تدريجي
ethnography	أثنوجرافيا
external criticism	نقد خارجي
focused interview	مقابلة مركزة
historical research	بحث تاريخي
human instrument	أداة بشرية
inductive analysis	تحليل استقرائي
internal criticism	نقد داخلي
naturalistic inquiry	بحث طبيعي
natural setting	محيط/ وضع طبيعي
nonparticipant observation	رصد/ ملاحظة غير المشارك
participant observation	رصد/ ملاحظة المشارك
primary source	مصدر أولي
qualitative research	بحث نوعي
quantitative research	بحث كمي
secondary source	مصدر ثانوي
triangulation	المسح المثلثي

تمارين EXERCISES

1- حدد أيا من العبارات التالية التي تميز الطريقة الكمية وتلك التي تميز الطريقة النوعية، حيث العبارة:

أ. تفترض أن أهداف وطرق العلوم الاجتماعية هي ذات أهداف وطرق العلوم الأخرى.

ب. ترى أن السلوك الإنساني محكوم دوماً بالإطار الذي يحدث فيه.

ج. تستخدم التفسيرات الافتراضية - الاستنتاجية.

د. تستخدم المسح المثلثي لتشييد الثقة.

2- كيف تختلف دراسات الحالة عن تجارب الفرد الواحد؟

3- يذكر عالم نفس بعض الأحداث في حشد لمجموعة توقعت في ليلة معينة أن كوكب الأرض سيختفي. ويتكون تقريره من ملاحظات فعلية لهذه المجموعة. فحسب أي من أنماط البحث التالية يمكن تصنيف الاستقصاء/ البحث؟

أ. بحث نوعي
ب. بحث تجريبي
ج. بحث تاريخي
د. بحث علّي مقارن

4- حدد الطرق التي يستخدمها الباحثون النوعيون لتشييد الثقة.

5- لدى باحث رسالة تصف التربية في أوغندا عام 1977. ويفترض أن الرئيس عيدي أمين كتبها.

أ. ما السؤال الذي سيطرح في النقد الخارجي؟
ب. ما السؤال الذي سيطرح في النقد الداخلي؟

6- متى تعتبر الوثيقة التاريخية ثانوية؟

7- ما هي مزايا ومساوئ البحث التاريخي مقارنة بأنواع البحث الأخرى؟

8- ميز بين ملاحظة المشارك وغير المشارك: أعط مثالا عن كل واحدة.

9- قارن دور الفرضية في البحث الكمي والنوعي.

10- أعط مثالا عن ملاحظة المشارك المتنكّر.

11- ما هي الملحوظات الميدانية؟

12- افترض انك أردت تحديد المدى الذي تناقش فيه كتب الدراسات الاجتماعية في المدارس الابتدائية إنجازات الأمريكان الزنوج. كيف ستمضي في الحصول على البيانات بشان هذه القضية ؟

13- أي مما يلي هي مشكلة بحثية نوعية ؟

أ. ما هي أساليب الإدارة الصفية التي يستخدمها مدرسو المدارس الثانوية في داخل المدن؟
ب. كيف تقارن رواتب مدرسي المدارس الثانوية داخل المدن مع رواتب مدرسي المدارس الثانوية الريفية أو الواقعة على الأطراف؟
ج. ما العلاقة بين سنوات الخبرة التدريسية والرضى عن العمل ؟

الأجوبة ANSWERS

1- أ. كمي
ب. نوعي
ج. كمي
د. نوعي

2- تركز تجارب الفرد الواحد على سلوك واحد أو عدد محدود جدا من أنواع السلوك. ويدخل الباحث معالجة محدودة بهدف دراسة تأثير هذا المعالجة على الأفراد. وتحاول دراسة الحالة وصف مجال سلوك الفرد الكلي كما يحدث في وضع طبيعي. ويلاحظ الباحث سلوك الفرد بالنسبة لتأثير في البيئة المادية والاجتماعية والنفسية.

3- أ

4- الالتزام المطول في الموقع، الملاحظة المتواصلة، المسح المثلثي- استخدام الملاحظات المضاعفة، مصادر البيانات، والطرق، واستخلاص النتائج بشكل دوري، والحفاظ على آثار التدقيق.

5- أ. هل كتب الرسالة عيدي أمين فعلاً؟
ب. هل تصف التربية بدقة في أوغندا عام 1977؟

6- تعتبر الوثيقة ثانوية إن تتدخل ذهن غير المراقب بين الحادث والوثيقة.

7- تكمن إحدى فوائد البحث التاريخي في عدم احتمال تأثير الباحث أو التفاعل التجريبي على تفسير النتائج، فبوسع المنظور التاريخي أن يتعامل مع القضايا والمواقف السابقة التي تتعذر معالجتها تجريبيا. وتكمن السيئة الرئيسية في الافتقار إلى الضبط التجريبي الذي يجعل التفسير الذي لا لبس فيه للبيانات والتعميمات أمرا صعباً. وثمة أيضاً، احتمال جمع معلومات غير كافية أو غير دقيقة بما يجعلها غير قابلة للإثبات.

8- في ملاحظة المشارك، يساهم الباحث بفعالية في المجموعة المدروسة. وفي ملاحظة غير المشارك، لا يساهم الباحث في أية فعالية، إنما يعمل فقط بوضعه ملاحظاً. تختلف الأمثلة.

9- الباحثون النوعيون لا يصوغون الفرضيات عادة قبل البحث ثم يقومون باختبارها بعدئذ. وتصاغ الفرضيات استقرائيا مع استمرار البحث. أما الباحثون الكميون، من ناحية أخرى، فإنهم يصوغون فرضية في بداية البحث، ثم يتابعون بشكل استنتاجي تحديد ما ينبغي ملاحظته إن كانت الفرضية صحيحة، ثم يجرون بعدئذ الملاحظات المطلوبة لاختبار الفرضية.

10- قد يتخذ الباحث عملا كنادل في نادٍ مجاور بغية دراسة هذه الوحدة الاجتماعية الخاصة.

11- الملحوظات الميدانية هي ملحوظات مسجلة في الميدان من قبل باحث يجري دراسة أثنوجرافية. وهي مكونة مما يراه أو يسمعه الباحث أو من تفسيرات الملاحظات.

12- سيكون تحليل المحتوى/ المضمون تصميماً بحثياً مناسباً. فالمرء يحصل على عينة من كتب الدراسات الاجتماعية الواسعة الاستخدام، ويشرع، بعدئذ، بالتصنيف ويمضي في تحقيق الكتب وإحصاء عدد المرات التي تذكر فيها إنجازات الأمريكان الزنوج ضمن كل صنف. وقد تكون هذه الأصناف ذكوراً أو إناثاً أو قد تنطوي على الحقول التي تمت فيها الإنجازات كالموسيقى والعلوم والآداب وما إلى ذلك.

13- أ

REFERENCES المراجع

Allen, A.M., Allen, D.N., and Sigler, C. (1993). Changes in sex-role stereotyping in Caldecott Medal award picture books 1938-1988. *Journal of Research in Childhood Education*, 7, 67-73.

Bogdan, R.C., and Biklen, S.K. (1992). *Qualitative research in education (2d ed.). Boston: Allyn & Bacon.*

Chilcott, J.H. (1992). Some contributions of anthropology to professional education. *Educational Researcher;* 21(2), 31-35.

Emerson, R.M. (Ed.). (1988). *Contemporary field research.* Prospect Heights, IL: Waveland.

Ceertz, C. (1980). Blurred genres: The refiguration of social thought. *American Scholar;* 49, 165-179.

Coetz, J.P., and LeCompte, M.D. (1984). *Ethnography and qualitative design in educational research.* New York: Academic Press.

Cuba, E.C., and Lincoln, Y.S. (1981). *Effective evaluation.* San Francisco: Jossey-Bass.

Hartshome, H., May, M. A., and Shuttleworth, F. K. (1928). *Studies in the organization of character.* New York: Macmillan.

Itard, J.C. (1962). *The wild boy of Aveyron.* (C. Humphrey & M. Humphrey, Trans.). New York: Appleton.

Janesick, V.J. (1994). The dance of qualitative research design: Metaphor, methodolatry, and meaning. In N. Denzin and Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 209-219). Thousand Oaks, CA: Sage Publications.

Kagan, D.M., Dennis, M.B., Igou, M., and Moore, P. (1993). The experience of being a teacher in residence. *American Educational Research Journal,* 30(2), 426-443.

Kraft, C.L. (1991). What makes a successful black student on a predominantly white campus? *American Educational Research Journal,* 28(2), 423-443.

Lancy, D.F. (1993). *Qualitative research in education. An introduction to the major traditions.* White Plains, NY: Longman.

Lincoln, Y.S., and Cuba, E.C. (1985). *Naturalistic inquiry.* Beverly Hills, CA: Sage Publications.

McCutcheon, C. (1981). On the interpretation of classroom observations. *Educational Researcher;* 10, 5-10.

Mischler, E.C. (1979). Meaning in context: Is there any other kind? *Harvard Educational Review,* 49, 2-10.

Peshkin, A. (1982). *The imperfect union: School consolidation and community conflict.* Chicago: University of Chicago Press.

Peshkin, A. (1986). *Gods choice: The total world of a fundamentalist Christian school.* Chicago: University of Chicago Press.

Peshkin, A. (1991). *The color of strangers, the color of friends: The play of ethnicity in school and community.* Chicago: University of Chicago Press.

Schwandt, T.A., and Halpern, E.S. (1988). *Linking auditing and metaevaluation.* Beverly Hills, CA: Sage Publications.

Spradley, J.P. (1980). *Participant observation* New York. Holt, Rinehart and Winston.

Strauss, A.L. (1987). *Qualitative analysis for social scientists.* Cambridge, UK: Cambridge University Press.



إيصال البحث

COMMUNICATING RESEARCH

* الخطوط العريضة لكتابة مقترحات البحوث.
* تحليل وتفسير وكتابة النتائج.

الفصل الرابع عشر

الخطوط العريضة لكتابة مقترحات البحوث
Guidelines For Writing Research Proposals

INSTRUCTIONAL OBJECTIVES أهداف تعليمية

بعد دراسة هذا الفصل سيكون بوسع الطالب أن:

1- يصف مكونات مقترح البحث وسبب أهميتها.

2- يحدد نقاط الضعف الشائعة/ العامة في مقترحات البحوث.

3- ينقد مقترحات البحوث.

4- يختار الإحصاءات المناسبة للأسئلة المطروحة والبيانات المعنية.

5- يصف الاعتبارات الأخلاقية والقانونية في البحث.

6- يصف دور مجلس المراجعة المؤسسي.

7- يكتب مقترح بحث يلبي هذه الخطوط العريضة.

سيحتاج الباحث في أغلب الحالات إلى عرض مشاريعه في صيغة مكتوبة منظمة بمرحلتين: 1) المرحلة الأولية التي تتطلب الإعداد لمقترح البحث 2) المرحلة النهائية وهي التقرير النهائي لنتائج البحث. ويعالج هذا الفصل المرحلة الأولية بينما يناقش الفصل 15 المرحلة النهائية.

كتابة مقترح البحث
WRITING A RESEARCH PROPOSAL

قد تكون كتابة مقترح البحث هي الخطوة الأهم والأكثر إثارة في العملية البحثية. ففي هذه المرحلة يتبلور المشروع برمته في صيغة ملموسة. وفي المقترح يبين الباحثون أنهم يعرفون ما يبحثون عنه وكيفية إدراكه كما يوضحون أسباب جدارة البحث. وتتم ترجمة تطلعات الباحث واستبصاره في خطط تدريجية لاكتشاف معرفة جديدة. وقد يكون التصميم/ الصيغة هو المخطط غير الرسمي نسبياً الذي يقدمه الطالب لتلبية متطلبات مقرر بحث أو يكون مقترحاً نظامياً لأطروحة يقدم إلى لجنة أو طلب تمويل إلى مؤسسة أو جهة حكومية.

ويحتوي المخطط المقترح التالي، لكتابة مقترح البحث، الخطوات الأساسية لصياغة واقتراح دراسة بحثية:

1- المقدمة
   أ. صياغة المشكلة
   ب. مراجعة المصادر
   ج. الأسئلة و / أو الفرضيات

2- طريقة / منهجية البحث
   أ. الأفراد
   ب. الأدوات
   ج. الإجراءات

3- تحليل البيانات
   أ. تنظيم البيانات
   ب. الإجراءات الإحصائية

4- أهمية الدراسة
   أ. المضامين
   ب. التطبيقات

5- الجدول الزمني والميزانية
   أ. الجدول الزمني
   ب. الميزانية

ومع أنه ليس من الضروري اتباع هذا المخطط بحذافيره، إلا أنه لابد من عرض دليل مفيد لكتابة أي مقترح لأن جميع الجوانب المدرجة هنا ينبغي لها أن تؤخذ بنظر الاعتبار.

المقدمة Introduction

الجزء المهم من مقترح البحث هو المقدمة. ففي الفقرة الأولى ينبغي أن يبيّن المؤلف مشكلة البحث بصورة واضحة دون لبس. وينبغي بعدئذ ربط المشكلة بالمعلومات المتيسرة في الحقل وبأهمية البحث والحاجة لتنفيذه. وبغض النظر عن مدى إحكام تصميم البحث ومدى جودة اختيار الإجراءات الإحصائية، فإنه ما لم تكتب المقدمة بشكل دقيق وواضح فإن الأجزاء الأخرى من المقترح سوف لن تحظى بالاعتبار الجاد. فليس من غير المعتاد رفض المقترحات لمجرد أن المقدمة ضعيفة دون اعتبار كبير لطريقة البحث المقترحة والتصميم الإحصائي. فمما يوصى به ضرورة أن يعد هذا الجزء بعناية وحذر وبهدف تعزيز اهتمام القارئ بالمشكلة.

وينبغي أن تحتوي مقدمـــة أي مقترح بحث 1) صياغة المشكلة، 2) مراجعة المصادر، 3) الأسئلة و/ أو الفرضيات.

صياغة المشكلة Statement of The Problem

يجب أن تتم الصياغة الواضحة والمباشرة للمشكلة بشكل مبكر جداً في المقدمة، ومثالياً، في بداية الفقرة الأولى حيث يعقبها وصف لخلفية المشكلة. ويجب أن يضم هذا الجزء من المقدمة إشارة موجزة لأهمية الدراسة، ومع ذلك فإنه ينبغي تجنب إغراء الترويج لأهمية الموضوع قبل عرضه. هناك اثنان من الأخطاء الشائعة ينبغي مراعاتهما وهما:

1- بدء المقدمة بعرض مدروس لخلفية المشكلة قبل الصياغة الواضحة للمشكلة ذاتها.

2- التركيز على تسويغ الدراسة في هذه النقطة، مع عدم التعرض لبيان المشكلة في النقاش، أو ذكرها بشكل غامض قرب النهاية فحسب.

خطأ شائع آخر هو الافتراض بان القارئ يعرف عن مضمون المشكلة بقدر ما يعرفه المؤلف. ينبغي ذكر عبارة المشكلة بحيث يمكن أن يفهمها شخص مطلع بشكل عام لكنه غير مطلع نسبيا على نطاق المشكلة.

وفي العادة ستتطور خلفية المشكلة في الجزء الخاص بالأدبيات ذات الصلة. ومن المفيد أحياناً، على أية حال، أن نذكر في صياغة المشكلة تلك الدراسات التي تقود إليها بشكل مباشر. وإذا برزت المشكلة من تجربة المؤلف، يمكن توضيح ذلك بشكل موجز في هذا الجزء.

ينبغي في مكان مناسب من هذا الجزء تحديد أية مصطلحات قد لا تكون مألوفة للقارئ، أو المصطلحات التي يعزو إليها المؤلف معان خاصة وذلك بذات الطريقة التي ستستخدم فيها في الدراسة. ويجب أن تنهي هذا الجزء بالقيود المحددة لمدى الدراسة والفرضية التي تُؤذن بها.

مراجعة الأدبيات Review of The Literature

يعرض المؤلف في هذا الجزء من الأدبيات ذات الصلة ما هو معروف حالياً عن المشكلة قيد الدرس، وبهذا يقدم الخلفية للأسئلة أو الفرضيات للدراسة المقترحة. اختر فقط الأدبيات ذات الصلة بأهداف الدراسة المقترحة. ولا ينبغي أن تكون الأدبيات المحددة شاملة، بل يجب أن تحتوي على الدراسات الأكثر صلة، وتوضح الوعي بالممارسات الحالية الواعدة.

في هذا الجزء لا يعرض مؤلف المقترح، فقط، كيف يقترح المضي مما هو معروف إلى المجهول، بل كذلك مدى ثبات قبضة المؤلف على الميدان وإدراكه للتطورات الأخيرة فيه. ولا يعني ذلك أن تكون المراجعة عملاً دالاً على ألمعية المؤلف. لذا ينبغي، فقط، تضمين الأدبيات المرتبطة بأهداف الدراسة.

وينبغي تنظيم الأدبيات حسب الموضوع. فالتنظيم حسب الموضوع يفيد في الإشارة إلى ما هو معروف لدى القارئ حول الجوانب المتنوعة للداسة. وهكذا توضع الصورة الكاملة لخلفية الدراسة معا خطوة بخطوة.

ثمة سقطة يجب تجنبها في الجزء الخاص بالأدبيات ذات الصلة، ألا وهو عرض سلسلة من الخلاصات، واحدة لكل فقرة. فهذا يقدم للجمهور قراءة مملة ويفقد فرصة وضع أساس ذي معنى للدراسة. فمن الأفضل بكثير التنظيم حسب الموضوع والإشارة إلى الكيفية التي ترتبط بها الدراسات بالأسئلة.

ولا حاجة إلى مناقشة جميع الدراسات ذات الصلة بالتفصيل. ففي مراجعة العديد من الدراسات المثيلة قد يصف المؤلف أهمها ثم يوضح ببساطة بان النتائج مؤكدة في دراسات مشابهة جاء ذكرها دون وصفها بالتفصيل. وغالباً ما يتصور الباحثون المبتدئون المتحمسون بأن دراستهم المقترحة فريدة وأنه لا يوجد هناك بحوث ذات صلة بها. وهذا نادراً ما يحدث. وغالباً ما يُبرز البحث المتعمق عدة بحوث لها صلة على الأقل ببعض الجوانب الخاصة بالدراسة المقترحة. وحتى إذا تعذر وجود بحوث في هذا المجال فإن هناك عادة أدبيات ذات طبيعة نظرية أو تأملية ينبغي تضمينها كجزء من خلفية الدراسة.

وينبغي على المؤلف، بالطبع، أن يضمّن نظريات ونتائج بحوث مخالفة للفرضية الموضحة وتلك التي تتفق معها.

ويجب أن ينتهي الجزء الخاص بالأدبيات ذات الصلة بمناقشة الاستنتاجات ومضامينها. هنا يتقاسم المؤلف الأفكار التي حصل عليها من مراجعة الأدبيات ويشير إلى الثغرات فيما هو معروف عن الموضوع، وبهذا يتوجه مباشرة إلى المسالة التي يقترح بحثها.

المسائل و/ أو الفرضيات Questions and/ or Hypotheses

إن المشكلة التي تمت صياغتها بطريقة عامة ينبغي الآن تحديدها. فإذا كان المشروع مسحاً، فينبغي صياغة المشكلة في سؤال – مثلا، ما هي نسبة المدرسين في ولاية أيوا ممن لديهم تثبيت في وظائفهم؟ وإذا صمم المشروع لاختبار نظرية معينة فينبغي، على أية حال صياغة المشكلة بصيغة الفرضية. ورغم أن الجواب على سؤال المسح قد يتخذ أي عدد من القيم، فإن الجواب في تجربة اختبار الفرضية هو دائما إما أن هناك دليلاً كاف لدعم الفرضية أو أن الدليل غير كاف.

تتحدد صياغة فرضية البحث بصورة نموذجية بمضامين الأدبيات ذات الصلة والمنطق الاستقرائي للدراسة. وتشير بعض المرجعيات إلى ضرورة صياغة الفرصة بالصورة الصفرية لأن الفرضية الصفرية هي المعنية في الاختبار الإحصائي. وعلى أية حال، نحن نشير إلى صياغة فرضية مقترح البحث في هذا الجزء في إطار العلاقات المتوقعة بين المتغيرات. وبهذه الطريقة يعطي المؤلف للقارئ مؤشراً واضحاً للقصد من الدراسة أكثر مما تحمله الفرضية الصفرية فيما لو تمت صياغتها في هذه النقطة. إن استخدام الفرضية البحثية في هذه المرحلة يسمح كذلك للباحث بناء منطق استنتاجي أساسي للدراسة. أما الفرضية الصفرية فيمكن إدخالها في الجزء الخاص بتحليل البيانات.

من المفضل بالنسبة للفرضية أن تصاغ بإيجاز بصورة إجرائية. وإذا لم يكن هذا ممكنا، فينبغي للفرضية المصاغة بشكل عام أن يتلوها التعريفات والشروط اللازمة لتعريفها بالصورة الإجرائية.

منهجية / طريقة البحث Methodology

في هذا الجزء من المقترح، يبين الباحث كيفية تنظيم الدراسة بحيث تتسنى الإجابة على سؤال البحث أو رصد العلاقات المفترضة، إن كانت مثل هذه العلاقات موجودة حقاً. وفي الفصول السابقة تم عرض تصميمات بحثية مناسبة لعدة أنواع من البحث. فينبغي على الباحث أن يختار من هذه التصميمات البحثية تجريبية كانت أو غير ذلك، تصميما يكون أفضل ما يلائم السؤال و/ أو الفرضية قيد الدرس. مثلا، إن تعين على المرء أن يقارن طريقتين لتدريس الكيمياء، فإنه بذلك يثير سؤالاً تجريبياً. وتتطلب مشكلة البحث هذه مجموعتين من الأفراد على الأقل: تجريبية وضابطة. وإن أراد أحدهم أن يستقصي أيضاً أثر التفاعل بين طرق تدريس الكيمياء ومتغير آخر – الاستعداد مثلاً – فالسؤال، الذي لا يزال تجريبياً يتطلب تصميماً أكثر تعقيداً من تصميم المجموعتين. فالمرء يحتاج، في هذه الحالة، إلى إقامة تصميم عاملي بأربعة مجموعات على الأقل لمثل هذه الدراسة.

وفي المنهجية كجزء من المقترح يشمل المؤلف جميع الخطوات التي ستتخذ لبحث المسالة قيد الدرس. إن إجراءات المعاينة المقترحة وطرق جمع البيانات والأدوات التي ستستخدم يتم وصفها.

إن إحدى الطرق المناسبة لعرض منهجية البحث تكون عن طريق تصنيف كل المعلومات فيما يخص التصميم: 1) الأفراد 2) الأدوات 3) الإجراءات حسبما يكون مناسباً.

الأفراد Subjects

وتكون الخطوة الأولى لتحديد الأفراد في دراسة ما هي وصف المجتمع الإحصائي المعني: فهل الدراسة معنية بالطلبة الجدد في الكلية أو الأطفال بسن السادسة ممن يعانون عسرة في القراءة، أو مديري المدارس الابتدائية، وما إلى ذلك؟ عندئذ يصف المؤلف/ الباحث إجراء سحب العينة من المجتمع الإحصائي. فإذا لم يكن الاختيار العشوائي ممكناً، فينبغي إيضاح سبب اتخاذ إجراء خاص لاختيار العينة ومدى تشابه العينة مع المجتمع الإحصائي المعني. إن وصفاً دقيقاً للأفراد يمكن أن يساعد القارئ للمقترح أن يحدد ما إذا كان بالإمكان، حسب وجهة نظر القارئ، تعميم نتائج الدراسة إلى المدى المطلوب.

الأدوات Instruments

إن هدف المشروع البحثي هو دراسة العلاقات بين المفاهيم البنائية. وعلى أية حال، فإنه نظراً لاستحالة قياس المفاهيم البنائية مباشرة، لابد لنا من اختيار أو تطوير مؤشرات ستقربها قدر الإمكان. وإذا توفرت أداة مسبقاً، فعلى المقترح أن يتضمن دليلاً على ثباتها وصدقها لأجل الدراسة. وفي الحالات التي ينبغي على الباحث بناء الأدوات، فإن من الضروري إيجاز الإجراء الذي يراد اتباعه في تطويرها. وينبغي على هذا المخطط أن يتضمن الخطوات التي ستتخذ للحصول على بيانات الصدق والثبات لهذه الأدوات. فإن أدى وصف الثبات والصدق إلى تفاصيل كثيرة قد تعيق استمرارية المقترح، فإن من المفضل تضمين هذه المادة في ملحق بدلا من النص.

الإجراءات Procedures

في الجزء الخاص بالإجراءات، يصف المؤلف الطريقة التي ستقام بها التجربة بحيث يمكن رصد العلاقات المفترضة، إن وجدت مثل هذه العلاقات بالفعل. وفي الحقيقة فإن الباحث يقول: "إذا كانت هذه الفرضية صحيحة، فعندئذ سترصد هذه النتائج". وبتصميم الدراسة كعملية تسمح بملاحظة العلاقات المفترضة، فإن الباحث يضع أساسا للدراسة.

إن الوصف الدقيق لإجراءات الدراسة متطلب أساسي لأي مقترح بحث. ففي البحث المسحي تكون كتابة هذا الجزء بسيطة نسبياً لأن الإجراء ينطوي فقط على إرسال الاستبيان لغرض ملئه أو إجراء المقابلات. ومع ذلك، فإنه ينبغي إدراج وتوضيح كل الخطوات – أي إعداد الاستبيان أو جدول المقابلة، تدريب المقابلين وإعطائهم التوجيهات حول كيفية التقرب إلى الأفراد وكيفية إنجاز المقابلة.

وقد تكون الإجراءات أكثر تعقيدا في البحث التجريبي. ففي هذا الجزء، ينبغي على المؤلف أن يدرج المجموعات، ويحدد التفعيلات المخططة لكل منها، خطوة بخطوة، ويربط كل معالجة بالأسئلة والفرضيات المقترحة. وينبغي تصميم هذه الخطوات بصورة كاملة بصيغة إجرائية. وبإقامة إجراءات التجربة على الفرضيات يسهّل المؤلف الطريق للتفسير المباشر والواضح للنتائج. وينبغي أن نأخذ بنظر الاعتبار احتمال المتغيرات المربكة – وهي تلك المتغيرات التي تعزى إلى فروقات درجات المعيار التي لا تكون جزءاً من المتغير المستقل. وينبغي أن تُحدد في الجزء الخاص بالإجراءات، الكيفية التي يقترح بها المؤلف/ الباحث ضبط هذه المتغيرات. مثلاً، قد يضبط المرء الفروقات بين الطلبة والمعلمين عن طريق تخصيص الطلبة والمدرسين عشوائيا للمجموعات الضابطة والتجريبية. وهنا يجب جعل الوقت المصروف والموقع المادي والتسهيلات متكافئة. وينبغي أن يتضمن هذا الجزء جميع الخطوات التي تعرف إجرائياً المعالجات التجريبية والضابطة.

أما الوثائق مثل مواد التدريس أو القراءة المخططة فلا حاجة لتضمينها في النص الأساسي للمقترح لان مناقشة مثل هذه التفاصيل تعيق في العادة استمرارية المقترح. ويوصى بان يضع المؤلف هذه الوثائق في ملحق يصفها باختصار ولكن بوضوح في جزء الإجراءات. وعلى أية حال فإن من المهم التوضيح في هذا الجزء أية فروقات في عرض المواد على المجموعات المختلفة المعنية في التجربة.

وبعد كتابة مسودة الجزء الخاص بالإجراءات، ينبغي قراءتها للتحقق من أن جميع الخطوات الضرورية للإجابة على كل سؤال واختبار كل فرضية تم وصفها. ويمكن التأكد من إكمال هذا الجزء بسؤال هو: هل يستطيع القارئ أن ينفذ هذا البحث باتباع الخطوات كما وصفت؟ فإذا كان بالإمكان الإجابة على هذا السؤال بالإيجاب يكون هذا الجزء كاملا.

تحليل البيانات Analysis of Data

يصف الجزء التالي من مقترح البحث طرق معالجة وعرض البيانات ويوجز الإجراءات الإحصائية المستخدمة. ويجب تجميع هذه المعلومات في أجزاء تغطي 1) تنظيم البيانات 2) الإجراءات الإحصائية.

تنظيم البيانات Data Organization

يمكن لعرض نتائج الدراسة البحثية أن يأخذ أشكالاً مختلفة استناداً إلى طريقة تنظيم الاستنتاجات. ومن الضروري التخطيط مسبقا لترتيب نتائج البحث في صيغة منظمة. ويتم ذلك بأفضل ما يكون عن طريق الإشارة إلى الأسئلة أو الفرضيات الخاصة بالدراسة. ان التخطيط المسبق لتنظيم وعرض البيانات يمكّن الباحث من تحديد ما إذا كانت المعلومات المجمعة ذات علاقة بالأسئلة البحثية. وإن من يتجاوزون هذه الخطوة غالباً ما يجدون أنهم قد بددوا وقتاً كثيراً وأموالاً في جمع شتات معلومات لا صلة لها.

تعد الجداول والأرقام والجداول وسائل أساسية لتنظيم وتلخيص جملة البيانات برمتها. فحينما يكون البحث في مرحلة التخطيط ينبغي على الباحث أن يكون قادراً على تصوير الكيفية التي سينظم بها الباحث البيانات ويقدمها في صيغة مجدولة. وفي هذه المرحلة، ينبغي على المرء أن يدرج أجزاء المعلومات المتيسرة لكل فرد، ويقرر كيفية عرض وتلخيص المعلومات، ثم يقرر ماهية الإجراءات الإحصائية التي سيستخدمها.

الإجراءات الإحصائية Statistical Procedures

إن تصميم الدراسة يحدد ما هي الأساليب الإحصائية التي ينبغي استخدامها وليس العكس. وبعبارة أخرى، يقرر الباحث أي تصميم سيسمح بملاحظة العلاقات المفترضة ثم يختار الإجراء الإحصائي الذي يناسب الأسئلة المطروحة وطبيعة البيانات المعنية. فالباحث لا يختار أولاً الإحصاءة التي تستهويه ثم يصمم الدراسة لتناسب تلك الإحصاءة (statistic).

لقد وصفت الإجراءات الإحصائية الأكثر شيوعاً في الفصول السابقة. وهي تتلخص للسهولة في الجدول 14.1 (الإحصاء الوصفي) والجدول 14.2 (الإحصاء الاستدلالي). إن الجدول 14.1 مصمم للمساعدة في تحديد المؤشرات التي قد تستخدم لوصف بيانات الدراسة بصيغة مجملة. ويتحدد الإجراء الإحصائي المناسب جزئياً بنوع سلّم القياس الذي يميز المتغير التابع. وعليه فإن الصفوف في الجدول تتحدّد بكونها فترية وترتيبية واسمية. أما الأعمدة (1) و (2) و (3) فتدرج الأهداف المتنوعة التي قد يخدمها الإحصاء الوصفي. أما الاستخدامات الأكثر شيوعا لهذه الإحصاءات فهي ما يلي:

1- تقديم مؤشر يصف مجموعة أو الفرق بين المجموعات (مقاييس النزعة المركزية).
2- تقديم مؤشر يصف التغير/ التباين في المجموعة أو الفروقات في تغير المجموعات (مقاييس التغير).
3- تحديد موقع الفرد في مجموعة (مؤشرات الموقع).
4- تقديم مؤشر يصف علاقة المتغيرات ضمن مجتمع إحصائي (مقاييس الارتباط)
5- وصف كيفية تقسيم مجموعة إلى مجموعات فرعية/ جزئية.
6- وصف التفاعل بين متغيرين أو أكثر حسب معيار ما (مقاييس التفاعل).

ويمكن تحديد الخلية المطلوبة عن طريق تحديد عنوان الصف والعمود المناسب لدراسة معينة. وتنقسم كل خلية ويتحدد المقطع المستخدم حسبما تكون الدراسة معنية بمجموعة أو أكثر. (تذكر أن المرء قد يختار إجراءاً لسلم قياس أدنى وليس العكس. مثلاً، قد يستخدم الوسيط أو المنوال لوصف البيانات الفترية ولكنه قد لا يستخدم الوسط الحسابي لوصف البيانات الترتيبية أو الاسمية).

الجدول 14.1: الإحصاء الوصفي

نوع سلم القياس للمتغير التابع	الغرض من الإحصاءة					
	(1) النـزعة المركزية		(2) التغير		(3) الموقع	
	مجموعة واحدة	أكثر من مجموعة	مجموعة واحدة	أكثر من مجموعة	مجموعة واحدة	أكثر من مجموعة
الفتري	الوسط	الفروق بين الأوساط	الانحراف المعياري للتباين	الفرق بين الانحرافات المعيارية للتباينات	الدرجة– Z أو أية درجة معيارية أخرى	الفرق بين الدرجة المعيارية للفرد في أكثر من توزيع
الترتيبي	الوسيط	الفرق بين الوسيطات	الانحراف الربيعي	الفرق بين الانحرافات الربيعية	الترتيب المئيني(*)	الفرق بين الدرجة المعيارية للفرد في أكثر من توزيع
الاسمي	المنوال	الفرق بين المنوالات / المناويل	المدى	الفرق بين أمداء (جمع مدى)	لقب أو تصنيف	لقب أو تصنيف

(*) لم يرد شرح الإحصاءة هذه في الكتاب ويمكن العودة إليها في العديد من كتب الإحصاء

تابع الجدول 14.1

نوع سلم القياس للمتغير التابع	الغرض من الإحصاءة					
	(4) الارتباط		(5) المجموعات الجزئية		(6) التفاعل	
	مجموعة واحدة	أكثر من مجموعة	مجموعة واحدة	أكثر من مجموعة	مجموعة واحدة	أكثر من مجموعة
الفتري	معامل بيرسون r	الفرق بين معاملات بيرسون للمتغيرات ذاتها في مجموعتين	---	---	الفرق بين أوساط الخلايا المرصودة والمتوقعة في ANOVA العاملي (التفاعل المرصود)	الفروق في التفاعل المرصود بين المجموعات
الترتيبي	معامل سبيرمان rho أو معامل كندول taw(*) أو W(*)	الفرق بين معاملات سبيرمان أو كندول	---	---	---	---
الاسمي	ارتباط ثنائي أصل أو ارتباط ثنائي، أو معامل phi	الفرق في الارتباطات للمتغيرات ذاتها في مجموعتين	النسبة أو النسبة المئوية	الفروق في النسب أو النسب المئوية	الفروق بين تكرارات الخلايا المرصودة والمتوقعة	الفروق في التفاعل المرصود بين المجموعات

(*) لم يرد شرح الإحصاءة هذه في الكتاب ويمكن العودة إليها في العديد من كتب الإحصاء

الجدول 14.2: الإحصاء الاستدلالي

نوع سلم القياس للمتغير التابع	الغرض من الإحصاءة					
	(1) النـزعة المركزية		(2) التغير		(3) الموقع	
	مجموعة واحدة	أكثر من مجموعة	مجموعة واحدة	أكثر من مجموعة	مجموعة واحدة	أكثر من مجموعة
الفتري	الخطأ المعياري للوسط	اختبار – t أو ANOVA ذي الطريق الواحد	------	اختبار بارتلت(*) أو اختبار –t لتجانس التباين(*)، إحصاء F-max(*)	الخطأ المعياري للقياس	الخطأ المعياري لدرجات الفرق (*)
الترتيبي	الخطأ المعياري للوسيط (*)	اختبار الوسيط ، اختبار الإشارة (*)، اختبار والس لـ ANOVA ذي الطريق الواحد(*) ، اختبار فردمان(*)	------	------	------	------
الاسمي	------	------	------	------	------	------

(*) لم يرد شرح الإحصاءة هذه في الكتاب ويمكن العودة إليها في العديد من كتب الإحصاء

تابع الجدول 14.2

نوع سلم القياس للمتغير التابع	الغرض من الإحصاءة					
	(4) الارتباط		(5) المجموعات الجزئية		(6) التفاعل	
	مجموعة واحدة	أكثر من مجموعة	مجموعة واحدة	أكثر من مجموعة	مجموعة واحدة	أكثر من مجموعة
الفتري	اختبار–t أو تحويل z لفشر أو اختبار F– للخطية(*)	اختبار –t أو تحويل z لفشر(*)	------	-------	اختبار –F لـ ANOVA متعدد العوامل	اختبار –F لـ ANOVA متعدد العوامل
الترتيبي	اختبار سبيرمان rho أو taw(*) أو W(*) لكندول	------	------	-------	------	------
الاسمي	اختبار مربع كاي للدلالة	Q لكوتشران (*)	مربع كاي أو اختبار ذي الحدين	مربع كاي أو اختبار فشر الدقيق / المضبوط (*)	التحليل اللوغاريتمي الخطي	اختبار مربع كاي للتحليل اللوغاريتمي الخطي

(*) لم يرد شرح الإحصاءة هذه في الكتاب ويمكن العودة إليها في العديد من كتب الإحصاء

ولتحديد سلم القياس الذي سيستخدم للتعبير عن البيانات، ينبغي على الباحث دراسة مزايا كل من سلالم القياس الثلاثة. فالبيانات الفترية تعطي معلومات أكثر من البيانات الترتيبية، والترتيبية تعطي معلومات أكثر من الاسمية. ولإقامة الاستدلالات فإن الاختبارات الإحصائية للبيانات الفترية أكثر "قوة" من اختبارات البيانات الترتيبية، أي أن للمرء فرصة أكبر في رفض الفرضية لدى استخدام المقاييس الفترية مما عليه الحال في المقاييس الترتيبية. وبالطريقة ذاتها، فإن الاختبارات الترتيبية أكثر قوة من المقاييس الاسمية.

فمثلا إذا كانت لدينا بيانات فترية لمتغير تابع وأردنا مؤشراً لوصف الفرق بين المجموعات، فإن الجدول يحدد الفرق بين وسطين حسابيين باعتبار ذلك إحصاءة مناسبة. (وبوسعنا أن نختار استخدام الفرق بين الوسيطات إلا أن ذلك سيكون أقل قوة من الفرق بين الأوساط الحسابية).

وإذا كانت الدراسة استدلالية بطبيعتها، فإن البحث سيمضي في اختبار الدلالة الإحصائية للمؤشر المختار. إن الإحصاءات المناسبة لهذا الغرض مدرجة في الجدول 14.2. ففي مثالنا يكون الاختبار التائي أو اختبار أنوفا (ANOVA) ذي الطريق الواحد مناسبا.

تذكر أن اختيار الإجراء الإحصائي يقوم على أساس كونه مناسبا للإجابة على سؤال تنطوي عليه الدراسة. ولا يمكن الحصول على شي باستخدام إجراء معقد عندما يقوم إجراء بسيط بالشيء ذاته. فالإحصاء هو في خدمة البحث ولا يهيمن عليه.

في هذا الجزء من المقترح، ينبغي إعطاء وصف محدد لخطط تطبيق الأدوات وجمع البيانات. وينبغي لهذه الخطط أن تتضمن الجدول الزمني، وإجراءات استبدال الأفراد المفقودين خلال مسار التجربة، وخطط الموازنة لترتيب التأثيرات وغيرها من التفاصيل الضرورية. إنا غالباً ما نقول لطلابنا "تصور أنك قد حصلت على منحة رائعة لإجراء دراستك إلا أن شاحنة دهستك في اليوم التالي، فهل بوسع زميل لك ان يأخذ مقترحك ويجري الدراسة فعلا؟" فإذا كان بالإمكان الإجابة على هذا السؤال بالإيجاب، فإن الجزء الخاص بتحليل البيانات سيكون كاملاً.

أهمية الدراسة Significance of The Study

يفضل بعض الباحثين بيان أهمية المشكلة في مقدمة المقترح. إن ترك هذا الموضوع لجزء تال، على أية حال، سيعطي الفرصة لربطها بخلفية وتصميم الدراسة. ويعالج هذا الجزء على نحو أفضل في مرحلتين: 1) المضامين، 2) التطبيقات.

المضامين Implications

نظراً لأن هدف البحث هو زيادة المعرفة، فعلى مؤلف المقترح أن يبين كيف أن دراسته ستقوم بذلك عن طريقة مناقشة النتائج التي ستساهم بالنظرية والمعرفة في مجال محدد ترتبط به مسالة البحث، ومدى فائدة هذه النتائج في حل المسائل والإجابة على أسئلة في الحقل العام. وأخيراً،

ينبغي أن يبين المؤلف كيف أن نتائج الدراسة ستقدم أسساً لبحوث تالية في هذا المجال. إضافة إلى ذلك، قد يوضح المؤلف / الباحث كيف أن خبرته وإطلاعه الواسع المقرونين بتسهيلات وأهداف المؤسسة موضع الدراسة، سيضعه في مكان مناسب لحل المشكلة قيد الدرس.

التطبيقات Applications

ينبغي أن يكون المؤلف قادرا على إقناع القراء بالتطبيق الممكن للاستنتاجات في الممارسة التربوية. وينبغي أن يبين هذا النقاش كيفية ومدى استفادة الممارسين التربويين من النتائج بغية تحسين عملهم. ولإيجاد المدى الذي يكون فيه للدراسة تطبيق في الممارسة التربوية، قد يسأل الباحث: هل ستغير نتائج دراستي أي شي في ميدان عملي؟ وهل ستساعد نتائجي المدرسين، ومرشدي المدارس، والمدراء، وغيرهم من التربويين في تحسين عملهم؟

ويحمل هذا الجانب وزناً كبيراً في جذب الكثير من الأموال المخصصة للبحث لغرض تنفيذ الدراسة. وتقيّم العديد من المؤسسات المقترحات البحثية على أساس التطبيق المحتمل في الممارسة.

الجدول الزمني والميزانية The Schedule and Budget

ينبغي التخطيط لجميع البحوث فيما يخص إمكانية تنفيذ العمل. وينبغي أن ينتهي المقترح بعرض 1) جدول زمني و 2) ميزانية.

الجدول الزمني Time Schedule

يحتاج الباحث إلى تحديد سياق المراحل الضرورية لإكمال البحث المقترح مع تقدير الوقت المطلوب لإكمال كل مرحلة. فتقسيم المشروع إلى مراحل ممكنة التنفيذ مع تحديد تاريخ لإكمال كل مرحلة، يساعد على تنظيم المشروع، ويمكّن الباحث من تقدير الجهد الذي سيحتاجه الباحث لإكمال المشروع برمته، وتامين فرص التقييم الدوري لتطوير المشروع.

الميزانية Budget

بمراجعة الأجزاء السابقة للمقترح، يحتاج الباحث الآن إلى تحديد الموارد الضرورية لتنفيذ البحث المقترح وإجراء التقديرات الدقيقة لتكاليف أمور مثل التجهيزات وتكاليف السفر، والاستنساخ، وخدمات الكومبيوتر، والعاملين وأجور الاستشارة. وإن كان الطلب يعتمد على تمويل خارجي، فإن لأغلب المؤسسات شخص، مثل المسؤول عن العقود، يساعد في تخطيط الميزانية لتناسب مستلزمات وكالة التمويل المتوقعة.

نقـــد المقتــــرح
CRITIQUING THE PROPOSAL

بعد إكمال تخطيط البحث، ينبغي على الباحث/ المؤلف أن يراجعه مرة أخرى بدقة وبعين ناقدة. ومن المفيد أيضاً الاستعانة بزملاء يقرأون المقترح. وفي الغالب يمكن لشخص آخر أن يحدد نقاط الضعف أو الإغفال/ الحذف غير الواضحة لدى المؤلف.

لقد حدد (Smith, 1963) من خلال عمله مع اللجنة الاستشارية للبحوث في برنامج البحوث التعاونية (USOE) ست نقاط ضعف مشتركة وجدتها اللجنة في المقترحات المقدمة للتمويل.

1- **المشكلة تافهة**: إن المشكلات ذات الأهمية السطحية للتربويين أو التي تبدي احتمالا قليلا بإضافة إلى المعرفة في التربية لا تعتبر مشكلات تستحق الدعم. ويعرض سميث مثالاً عن خطة لدراسة برنامج تربوي للكبار بالنسبة لأمناء المكتبة، ويعلق بقوله: "إنها ليست فقط ذات قيمة سطحية للتربية وذات اتجاه نحو العمل وليس البحث، بل كذلك إنها مشكلة ذات أهمية قليلة" (p. 282).

2- **المشكلة غير محددة**: يعتبر الإيضاح الكلاسيكي الذي قدمه كود (Good & Skates, 1954) مثالاً رائعاً حول هذا الضعف. فقد أوضحت رسالة كتبها طالب خريج إلى مسؤول التربية في الاسكا أن الطالب كان قد اختار موضوعا لأطروحته: "تدريس الإنجليزية كما يتضح في الدورات الدراسية لشعوب العالم الناطقة بالإنجليزية". وفي الفقرة الثانية يسال الموظف المسؤول: "هل تعرف بعض الكتب المهمة حول الاسكا، وتاريخها ومشكلاتها الاقتصادية، وتجارتها، وصادراتها، ووارداتها، وعلاقاتها الإنسانية ودياناتها، الخ.

ولتقديم مقترح ممكن، يجب على الباحث أن يركز الدراسة. ولا يعني ذلك أنه لا ينبغي على الدراسة أن تنطوي على عدد من المتغيرات ذات العلاقة. فعلى الباحث أن يتصدى لجوانب المشكلة التي يمكن معالجتها بصورة معقولة في دراسة واحدة. ويمكن للدراسة أن تشتمل على مجموعة من المتغيرات ذات الصلة، وهذا ما يجب أن يكون غالباً، غير أنه ينبغي تجنب الجهود غير العملية أو المبالغ فيها.

3- **صياغة الأهداف والفرضيات أو الأسئلة بشكل واسع**: يقع كاتبو المقترحات في الغالب تحت طائلة الإغراء في صياغة أهدافهم أو فرضياتهم أو أسئلتهم بشكل تعميمات واسعة وجارفة. وفي مثل هذه الحالات، يجد المرء عند قراءته للجزء الخاص بالإجراءات أن الدراسة الفعلية المخطط لها غير قادرة على تلبية الأهداف الكبرى المعلنة. وهذه الأهداف أو الفرضيات قد تصاغ أحياناً بشكل واسع وعام بحيث يتعين على المرء أن يعود إلى الإجراءات لاكتشاف

عما تدور حوله الدراسة. ويبدو من الواضح ضرورة أن تتفق الأهداف والإجراءات، ومع ذلك فإن لجنة البحوث الاستشارية وجدت إخفاق العديد من المقترحات في تلبية هذا المتطلب الأساسي.

4- **الإجراءات تفتقر إلى التفصيل**: يشير سميث إلى أن "الباحث الذي يحذف أكثر مما يضمّن، لا يتعين عليه أن يتوقع من اللجنة أن تقرا تفاصيل الإجراءات بين السطور. فاللجنة تفتقر إلى الرغبة البصيرة للقيام بذلك. "تذكر ضرورة أن تكون الإجراءات كاملة بما يكفي لتكرار / لاستنساخ الدراسة.

5- **استخدام تصميم بسيط لبحث مشكلة معقدة**: ينبغي أن يكون تصميم الدراسة مناسبة للمشكلة. فالمقارنة البسيطة لوسطي مجموعتين مناسبة حين ينطوي ذلك على متغير واحد. أما الدراسات الأكثر تعقيداً فإنها تتطلب تصميمات أكثر تعقيداً.

6- **عدم الاهتمام بالمتغيرات ذات الصلة أو إهمالها بخفة**: إن الإخفاق في اخذ المتغيرات الدخيلة ذات الصلة بنظر الاعتبار هو خطأ جسيم في مقترح البحث. فيتعين على الباحث أن يبين انه مدرك لمثل هذه المتغيرات ويوضح كيفية معالجتها في تصميم الدراسة.

أهمية إكمال المقترح قبل جمع البيانات

THE IMPORTANCE OF COMPLETING THE PROPOSAL BEFORE COLLECTING DATA

في البحث الكمي، يبين المقترح الواضح الكامل أن الباحث المأمول مستعد فعلاً لبدء الدراسة. فهو يكشف أن الباحث يعرف ما يفعله ولماذا يفعله وكيف يفعله. أما الباحث الكمي الذي ليس بوسعه تقديم مقترح كامل ومتماسك فإنه غير مستعد للمضي نحو مرحلة جمع البيانات الخاصة بالمشروع. ويميل الباحثون المبتدئون غالبا إلى القول: "دعني أجمع البيانات الآن وأقرر ما يمكن فعله فيما بعد". وقد يبدو أن جمع البيانات وكتابة المقترح في الوقت ذاته هو إجراء يوفر الوقت، إلا أن هذه الحالة نادرة. فقد أنفقت ساعات عمل كثيرة وآلاف الدولارات بتلك الطريقة. فما لم تتم صياغة المقترح فإنه لا يمكن للمرء أن يكون متأكدا بالضبط عما يحتاجه من بيانات أو ما هي الطريقة المثلى لمعالجة المعلومات في ضوء هدف الدراسة. وينبغي على من يعملون دون تحديد مواعيد للإنجاز أن يضعوا لأنفسهم تاريخا لإنهاء المقترح سابقاً لتاريخ إكمال المشروع برمته.

لا تحدد البيانات المراد جمعها في البحث النوعي مسبقا غير أن تصميماً يبرز مع تقدم الدراسة. ففي المقترح النوعي، يحتاج الباحث إلى أن يوضح مسبقا هدف البحث، والأفراد والوضع المراد دراسته، والمراحل الأولية للبحث.

الاعتبارات الأخلاقية والقانونية
ETHICAL AND LEGAL CONSIDERATIONS

يعتبر الالتزام الصارم بالمعايير الأخلاقية في تخطيط البحث وتنفيذه أمراً بالغ الأهمية. فعلى الباحثين التزامات إزاء الأفراد ومهنتهم. وكما أشرنا في المقاييس الأخلاقية المقترحة لجمعية البحوث التربوية الأمريكية (AERA)، فإننا كباحثين تربويين:

> **... يتعين علينا الحفاظ على سلامة بحثنا ومجتمعنا البحثي وجميع الذين لدينا معهم علاقات مهنية. ينبغي أن نعاهد أنفسنا على القيام بذلك بالحفاظ على كفاءتنا وكفاءة الناس الذين نزج بهم في الميدان عن طريق تقييمنا المستمر لبحثنا من أجل وفائه الأخلاقي والعلمي، وعن طريق إدارة علاقاتنا الداخلية والخارجية حسب أعلى المعايير الأخلاقية (AERA, 1991).**

الالتزام نحو الأفراد Obligation to Subjects

يجب على المرء لدى استخدام أفراد في البحث أن يحترم حقوقهم وكرامتهم وخصوصيتهم وحساسياتهم. وتتحدد عشرة مظاهر لالتزامات المرء نحو الأفراد في المعايير المقترحة لجمعية البحوث التربوية الأمريكية (AERA, 1991):

1- أن للمشاركين أو أولياء أمورهم في الدراسة البحثية الحق في الاطلاع على المخاطر المحتملة في البحث وعن العواقب المحتملة للمشاركين وعليهم أن يعطوا موافقتهم الواضحة قبيل المشاركة بالبحث. وينبغي على الباحثين التربويين إبلاغ من ينوب عنهم والمشاركين (وأولياء أمورهم) أو من ينوب عنهم من المؤسسات، بأهداف البحث وتحديث معلوماتهم عن أية تغيرات مهمة في برنامج البحث.

2- ينبغي أن تكون النزاهة سمة العلاقة بين الباحثين والمشاركين، وممثلي المؤسسات. فلا يتم تشجيع الخداع الذي ينبغي استخدامه فقط عند الضرورة في الدراسات العلمية كما ينبغي تقليصه. ويجب أن يوضح الباحثون للمشاركين وممثلي المؤسسات أسباب الخداع بعد الدراسة.

3- ينبغي أن يكون الباحثون التربويون حساسين لأية سياسات مؤسساتية قائمة خاصة أو إرشادات لإجراء البحث.

4- للمشاركين الحق في الانسحاب من الدراسة في أي وقت.

5- ينبغي أن يمارس الباحثون التربويون الحذر لضمان عدم استغلال أي منفعة شخصية لمجتمعات البحث أو الأوضاع المؤسسية للبحث. ولا ينبغي للباحثين التربويين

استخدام تأثيرهم على التابعين لهم، أو الطلبة أو الآخرين وذلك لإجبارهم على المشاركة في البحث.

6- على الباحثين مسؤولية مراعاة الفروقات الثقافية والدينية، والجنس، والفروق المهمة ضمن مجتمع البحث في ميدان التخطيط، والتنفيذ، وكتابة تقرير البحث.

7- ينبغي على الباحثين مراعاة وتقليص استخدام الأساليب البحثية التي قد يكون لها عواقب اجتماعية سلبية، مثل القياسات الاجتماعية السلبية مع الصغار أو التدخلات التجريبية التي قد تحرم الطلبة من أجزاء مهمة من المنهج التدريسي المقرر.

8- ينبغي أن يكون الباحثون التربويون حساسين إزاء سلامة النشاطات المؤسسية القائمة، كما يجب أن ينبهوا ممثلي المؤسسات المناسبين عن التشويش المحتمل في مثل هذه النشاطات التي قد تنجم عن إجراء البحث.

9- ينبغي على الباحثين التربويين إيصال استنتاجاتهم وأهميتها العملية بلغة واضحة ومباشرة ومناسبة إلى المجتمعات المعنية وممثليها وغيرهم ممن لهم صلة.

10- للمشاركين ومساعدي الباحث الحق في طمس هويتهم. وينبغي احترام هذا الحق عندما لا يتم التوصل إلى فهم واضح على عكس ذلك. فالباحثون مسؤولون عن توخي الحذر المناسب لحماية سرية كل من المشاركين والبيانات. وينبغي لمن تتم دراستهم أن يدركوا إمكانيات تقنيات جمع البيانات المختلفة المستخدمة في البحث كي يتمكنوا من اتخاذ قرار مطلع حول مشاركتهم. وينبغي الإيضاح للمساعدين والمشاركين أنه رغم كل جهد تم للحفاظ على البحث فقد تتم المساومة حول إغفال الأسماء. وعلى الباحثين الثانويين احترام إغفال الأسماء والحفاظ عليها حسبما قرره الباحثون الرئيسيون.

الالتزام نحو المهنة Obligation to the Profession

الباحث مسؤول أيضاً أمام المعنيين بالبحث. إن أغلب الدراسات البحثية، في التربية كما في الميادين الأخرى، يتم نشرها في المجلات والدراسات العلمية، والكتب، ووسائل الإعلام الأخرى ويشار إليها ويستثمرها المحترفون في الحقل. ويعتبر الباحث ملزماً بصورة أخلاقية بتخطيط الدراسة بطريقة ما بحيث لا تؤدي النتائج إلى إعطاء معلومات مضللة. والأكثر من ذلك، فإن الباحث ملزم بذكر النتائج، أيا تكون، بصورة نزيهة ومضبوطة. فلا ينبغي عرض البحث بطريقة تؤدي إلى التضليل . فالتقرير الذي يضلل يعد خللاً فادحاً لمسؤولية الباحث إزاء المهنة.

إن ما يفزع هو السهولة في إجراء بحث يصبح خادعاً بشكل غير مقصود. كما أشار (Kromrey, 1993):

نحن كباحثين نكرس الوقت والطاقة لبسط نظرياتنا، إلى المدى الذي تصبح فيه نظرياتنا مثل أطفالنا. ونتيجة لذلك، نغدو مثل الأباء ذوي المنظار الضيق الذي لا يسمح لهم برؤية أي خلل في أطفالهم، فنقنع أنفسنا بأن البيانات التجريبية ستشهد لصالح نظرياتنا إن كان بوسعنا فقط ان نحللها "بصورة مناسبة". ومن هذا المنظور تصبح جهودنا البحثية موجهة نحو دعم نظرياتنا بدلا من اختبارها.

ففي بحث مسحي قد يصوغ أحدهم، بغير قصد، أسئلة بطريقة تمكنه من تحديد الأجوبة التي يبحث عنها. وبإجراء البحث قد يرتب أحدهم، بدون انتباه، الأشياء بحيث يكون للمجموعة التجريبية المزايا التي لا تمتلكها المجموعة الضابطة والتي هي ليست جزءاً من المتغير المستقل. وقبيل البدء بالبحث، يوصى بأن يراجع زميل موثوق بدقة خطة البحث بحثاً عن مصادر محتملة لنتائج مضللة.

فإذا أخفقت التحليلات الإحصائية في دعم أو نقض فرضية معينة، فإنه يكون مغرياً الافتراض أن تحليلاً إحصائياً مختلفاً قد "يظهر صحيحاً". من المهم اختيار التحليل الإحصائي، الأكثر ما يناسب، قبل إجراء الدراسة لنستمد الاستنتاجات من تلك الإحصائيات. ومن المهم أيضاً ذكر جميع نتائج التحليلات الإحصائية وليس فقط تلك التي تكون دالة إحصائياً. فمثلاً، إذا قام المرء باختبار خمسين فرضية ووجد أن اثنتين منها دالة إحصائياً، فإن ذكر الاثنتين فحسب لابد أن ينطوي على أن الأمر لا يحتمل أن يكون محض صدفة – فعندما تكون أثنتان دالتين إحصائيا من خمسين فثمة احتمال كبير في أن يعود ذلك إلى الصدفة.

إن من السهل جداً، باستخدام الحاسوب، البحث خلال بيانات المجموعات الجزئية للعينة التي تتأكد فيها فرضيات المرء رغم القبول بالفرضيات الصفرية لمجموع العينة. فمثلاً، قد يكتشف المرء ما يبحث عنه في مجموعة طالبات الصف السادس ممن يحملن ألقابا أسبانية. فمن المناسب التخطيط للتباينات الإحصائية بين المعالجات وبين أنواع الجنس، وبين المستوى الدراسي، وبين المجموعات العرقية قبل إجراء الدراسة. وعلى أية حال، إن نتائج التحليلات غير المخطط لها للمجموعات الجزئية يمكن أن تؤدي إلى نتائج مضللة.

الالتزامات القانونية Legal Obligations

إن التزامات الباحثين بأفراد البشر لا تندمج فقط في العبارات الأخلاقية للمؤسسات المهنية مثل الجمعية التربوية الأمريكية (1991) والجمعية النفسية الأمريكية (1992)، بل كذلك تندمج في الأنظمة الفدرالية.

فاستخدام الأفراد البشرية في البحوث الممولة بشكل رسمي قد خضع للتنظيم منذ عدة عقود من الزمن، إلا أنه في عام 1975 تم اتخاذ مجموعة قواعد فعلية أكثر صراحة لتشمل جميع البحوث التي تقوم بها أو تشرف عليها جميع المؤسسات التي تتلقى تمويلاً حكومياً لإجراء

البحوث. ونتيجة لهذه القواعد، فإن أغلب الكليات والجامعات والعديد من الأنظمة المدرسية قد أقامت مجالس رسمية للمراجعة (IRB) تدعو إلى إجراء مراجعة مسبقة لجميع الخطط البحثية التي تنطوي على استخدام البشر. وهناك ثلاثة مجالات رئيسية ذات اهتمام وهي: 1) حماية الأفراد من الأذى 2) احترام حقوق الأفراد في معرفة طبيعة وهدف الدراسة وحقهم في إعطاء أو سحب موافقتهم على المشاركة (حق الموافقة بعد الإطلاع) و 3) احترام خصوصية الأفراد .

تحدد القوانين الحالية، القوانين الفدرالية (الرسمية) الخاصة بالمشاريع التي تمول عن طريق منح من وزارة الصحة والخدمات الإنسانية وتعطي إعفاءات مطلقة واسعة من هذه القوانين. أما الإعفاءات الخمسة الكبرى المطلقة التي حددها (قانون الأنظمة الفدرالية) (1993) فهي كما يلي:

1- البحوث التي تجرى في بيئات تربوية عريقة أو مقبولة بشكل عام وتضم ممارسات تربوية اعتيادية مثل (أ) البحوث في استراتيجيات التدريس للتربية الاعتيادية والخاصة أو (ب) البحوث في الفاعلية أو المقارنة بين أساليب التدريس أو المناهج أو طرق إدارة الصف.

2- البحوث التي تنطوي على استخدام الاختبارات التربوية (الإدراكية/ المعرفية، التشخيصية، الاستعداد، الإنجاز/ التحصيل)، الإجراءات المسحية، المقابلات أو مراعاة السلوك العام إلا إذا: (أ) تم تسجل المعلومات المستحصلة بطريقة يمكن فيها تحديد الأفراد بصورة مباشرة من خلال معرفين لهم علاقة بالأفراد، و (ب) كان أي كشف لاستجابات الأفراد خارج نطاق البحث قد يضع الأفراد بشكل منطقي إزاء خطر إجرامي أو مسؤولية مدنية أو يلحق الضرر بموقفهم المالي أو المهني أو سمعتهم .

3- البحوث التي تعني باستخدام الاختبارات التربوية (إدراكية/ معرفية، تشخيصية، استعداد ، إنجاز/ تحصيل)، إجراءات مسحية، مقابلات، أو مراعاة السلوك العام غير المعني حسب الفقرة (ب) من هذا الجزء إذا:
(أ) تم اختيار الأفراد أو الموظفين المعنيين أو المرشحين لمنصب رسمي، أو (ب) تطلبت القوانين الفدرالية بدون استثناء الحفاظ على سرية المعلومات التي تحدد الأشخاص خلال البحث وما بعده.

4- البحوث التي تعنى بجمع أو دراسة البيانات الحالية. والوثائق، والسجلات، والعينات المرضية، أو العينات التشخيصية، إن توفرت هذه المصادر بشكل عام أو إذا قام الباحث بتسجيل المعلومات بطريقة لا يمكن فيها تحديد الأفراد مباشرة أو من خلال معرفين مرتبطين بالأفراد.

5- مشاريع البحوث والشواهد العملية التي تجريها وزارة/ دائرة أو وكالة، أو تجري

بموافقة رؤسائها، التي تكون مصممة لغرض الدراسة أو التقييم أو الاختبار.

(أ) برامج الصالح العام أو الخدمات، (ب) إجراءات الحصول على المنافع أو الخدمات حسب هذا البرنامج، (ج) التغيرات المحتملة أو البدائل لتلك البرامج أو الإجراءات، أو (د) التغيرات المحتملة في طرق أو مستويات الأجور للفوائد أو الخدمات حسب هذه البرامج.

وبينما تعفي هذه الأنظمة الجديدة الكثير من البحوث التربوية من الأنظمة الفدرالية المباشرة، إلا أن مجالس المراجعة الرسمية (IRBs) تبقى في مكانها، وقد تختار المؤسسات المضي مع الأنظمة التي تعد أكثر تقييداً في طبيعتها أو أوسع في مداها من الأنظمة الفدرالية. وقد كان هناك ضغط غير رسمي على المؤسسات في الاستمرار بالأنظمة الأوسع. وينبغي على الباحث أن يخطط لاستشارة مجلس المراجعة (IRB) الذي له علاقة به لتحديد القواعد والسياسات المؤسساتية.

حماية الأفراد من الأذى Protection of Subjects Form Harm

أما بالنسبة إلى البحوث التي تمولها وزارة الصحة والخدمات الإنسانية والتي تتضمن أناسا من خارج نطاق الأصناف المعفاة، فإن الأنظمة الرسمية تتطلب تحديدا فيما إذا كان الأفراد سيتعرضون "إلى خطر" عن طريق إجراءات البحث أم لا. فإذا كان بمقدور الباحث أن يثبت بأن المخاطر الاجتماعية والبدنية والعقلية التي ينطوي عليها مشروع البحث ليست أكبر من التي يواجهها الأفراد في "الحياة اليومية أو خلال أداء الفحوص أو الاختبارات البدنية أو النفسية " فلا يُعدّ الأفراد في موضع "خطر". فإن كانوا موضع "خطر" فيجب على الباحث أن يكون قادراً على بيان إلى أن (1) قيمة المعرفة التي يحتمل الحصول عليها تفوق الأذى المحتمل، (2) سيطلع جميع الأفراد على مخاطر الإجراءات وسوف يتم الحصول على موافقة طوعية وتحريرية من كل فرد، (3) ستتوفر الخدمات الطبية والأخرى المساندة للأفراد المشاركين في البحث. فإذا تمت تلبية هذه الشروط فإن المجلس (IRB) سوف يعطي الموافقة على المقترح، غالباً. وعادة، ليس هناك إجراء بالتماس ذلك من المجلس.

الموافقة المطّلعة Informed Consent

رغم أن الأنظمة الفدرالية تتطلب الآن موافقة رسمية من الفرد في مجال محدود من الحالات فقط، إلا إن هناك العديد من المجموعات غير الحكومية للأخلاقيات المهنية ترى وجوب اطلاع كل فرد على أهداف الإجراءات وإعطاء خيار حر تام بالمشاركة. يجب على الباحث أن يأخذ بنظر الاعتبار الرغبة باطلاع الأفراد على البحث والحصول على موافقة رسمية منهم. وتحظر القواعد المهنية للمعايير الأخلاقية فعلاً، إكراه الأفراد والتوضيحات المخادعة للهدف.

قد يشار، في بعض الحالات، إلى أن المعرفة المسبقة بأهداف الدراسة قد تعمل على تحيز النتائج. وبالنسبة لحالات كهذه، لا تزال هناك أهمية في الحفاظ على الطبيعة الطوعية للمشاركة حتى وإن لم تكن الأهداف معروضة بالكامل أو بشكل دقيق قبل الإجراءات. وفي الحالات التي يكون فيها من المهم عدم مناقشة الهدف قبل الإجراء، يجب إعطاء توضيح كامل لجميع الأفراد بعد إكمال الإجراءات.

إن الحصول على الموافقة المطّلعة من القاصرين والأشخاص الآخرين الذين لا يعتبرون قادرين على تمثيل أنفسهم يُعدّ مجالاً صعباً. وينصح الباحثون الذين يواجهون هذه المشكلة باستشارة رئيس IRB الذي يتبعونه (إن كانوا مؤسسة) أو المستشار القانوني لمؤسستهم أو منظمتهم. إن مكتب الحماية من مخاطر البحث في وزارة الصحة والخدمات الإنسانية سيكون مستعداً في الغالب لتقديم الاستشارة حول هذه القضايا، وقد يكون قادرا على توفير المعلومات الجارية حول القواعد القانونية الخاصة بالجماعات المتنوعة.

الحق في الخصوصية Right to Privacy

هناك اتفاق رسمي واسع بأن لجميع الأفراد حقا متأصلا بالخصوصية. فإذا لم يكن من الضروري جمع اسم الشخص والمعلومات المعرّفة الأخرى، فيوصى بعدم جمع المعلومات. وإذا كان من الضروري جمع البيانات لأجل المتابعة أو أهداف أخرى، فإنها من مسؤولية الباحث عندئذ القيام بالتخزين الآمن لتلك المعلومات وضبط الوصول إليها. وبشكل عام، يوصى بأن يكون للباحث الرئيسي والأعضاء الذين يجب أن يعرفوا بيانات الشخص المتعلقة بأهداف العمل الحق في الوصول إليها. فإذا تعين إعطاء الطلبة أو الزملاء فرصة للوصول إلى البيانات أو وضعها في الأرشيف، فإن من مسؤولية الباحث الرئيسي عندئذ إزالة جميع المعلومات الشخصية المعرّفة من هذه الأضابير.

أما بالنسبة للبحوث التي تمولها وزارة الصحة والخدمات الإنسانية، فإنه يجب على الباحث توفير ضمان التخزين الآمن والحفاظ على الأسماء والمعلومات المعرّفة الأخرى. فإذا جمع مشروع البحث أية معلومات جرمية شخصية أو مضرة اجتماعيا فإن هذا الالتزام سيكون أكثر أهمية. إن الإخفاق في الحفاظ على المعلومات من هذا النوع قد يؤدي إلى عقوبات حسب الأنظمة الفدرالية (نزع الأهلية ومنح الدعم اللاحقة) أو المقاضاة المدنية بسبب الأضرار الشخصية.

الخلاصة SUMMARY

يعد مقترح البحث خطة تدريجية لاكتشاف معرفة جديدة. ففي هذه المرحلة يتبلور إلهام وتفكير الباحث في شكل ملموس. وينبغي إدخال عدة أصناف من المعلومات في مقترح البحث.

إن الصياغة الواضحة للمشكلة، بما يصحبها من تعريفات غير غامضة للمصطلحات يجب

أن يتم في وقت مبكر في المقترح. ويجب أن يتبع ذلك مراجعة الأدبيات ذات الصلة. والمراجعة الجيدة للأدبيات تبين ما هو معروف حتى الآن حول المشكلة وتضع الأساس لصياغة الفرضيات بما يخص العلاقات بين المتغيرات قيد الدرس. إضافة إلى ذلك، ينبغي كتابة هذا الجزء بهدف وضع الأساس لتفسير النتائج.

وفي مناقشة المنهجية التي تلي الجزء الخاص بالمقدمة، فإنه ينبغي وصف طرق اختبار الأفراد، وطرق جمع البيانات، وإجراءات الرصد، وأساليب القياس بتفصيل كاف لكي يتمكن قارئ ما من إجراء البحث باتباع الخطوات المقترحة تماماً كما سيفعل كاتب المقترح الأصلي.

ويصف الجزء التالي من مقترح البحث الإجراءات التي ستستخدم لعرض البيانات، مثل الجداول، والأشكال، والرسوم، كما يقدم الأساليب الإحصائية التي ستستخدم لتحليل البيانات.

يتبع ذلك مناقشة الأهمية المحتملة للدراسة. وهنا ينبغي على الباحث أن يحاول بيان كيف أن النتائج ستزيد المعرفة وما تعنيه هذه النتائج للنظرية والبحث في الحقل قيد الدرس. إن مناقشة التطبيقات الخاصة بالاستنتاجات في الممارسة ستكون مفيدة للقراء الذين يرغبون في تقييم أهمية البحث المقترح.

أما الجزء الأخير من المقترح فيحتوي على الجدول الزمني والميزانية المقدرة للدراسة. وتعدّ هذه المعلومات مفيدة للقراء في بناء تقييم شامل للمقترح.

وثمة أهمية كبيرة في تخطيط البحث وهي مراعاة المعايير الأخلاقية. فينبغي أن يكون للأفراد حق الموافقة المطّلعة، ويجب حمايتهم من الأذى كما يجب احترام خصوصيتهم.

مفاهيم أساسية Key Concepts

categorical exemptions	الاستثناءات/ الإعفاءات المطلقة
Code of Federal Regulations	قانون الأنظمة الفدرالية
common faults in research proposals	الأخطاء الشائعة في مقترحات البحوث
ethical considerations in research	الاعتبارات الأخلاقية في البحث
informed consent, right to	حق الموافقة المطّلعة
ingredients of research proposals	عناصر مكونات مقترحات البحوث
institutional review board (IRB)	مجلس المراجعة المؤسساتي (IRB)
legal considerations in research	الاعتبارات القانونية في البحث
protection of subjects from harm	حماية الأفراد من الأذى
right to privacy	حق الخصوصية

تماريـــن EXERCISES

1- ما هي العناصر الأساسية لمقترح البحث؟

2- ما هو سبب أهمية كتابة المقدمة بدقة ووضوح؟

3- عند أي نقطة من المقترح، ينبغي عرض صياغة واضحة للمشكلة؟

4- اكتب الفرضيات التالية بطريقة تعرض كل المتغيرات إجرائياً:

أ. أطفال يتعلمون القراءة بالطريقة الكلية للغة يقرأون أفضل من الأطفال الذين يدرسون بالطريقة التقليدية.

ب. طلبة الثانوية ممن يحرزون درجات فوق الربيع الأعلى في اختبار XYZ للاستعداد الميكانيكي يؤدون افضل في دروس الميكانيك.

ج. درجات طلبة الصف السابع الأذكياء ممن درسوا حسب طريقة الرياضيات الحديثة لمدة سنة، في المجموعات الجزئية لاختبار التحصيل SRA، تزيد عن درجات طلبة الصف السابع الأذكياء الذين درسوا بموجب الطريقة التقليدية.

5- ما هي بعض المتغيرات المربكة التي قد تؤثر على فروقات أوساط درجات التحصيل بين صفوف المرحلة؟ وكيف تضبط على هذه المتغيرات في إجراءاتك المقترحة؟

6- ما هي الإحصاءات المناسبة لقياس الارتباط إن كان سلم قياس المتغيرات أسمياً؟

7- ما هو مقياس النزعة المركزية المناسب للبيانات الفترية؟

8- ما أنواع البيانات التي يكون فيها (ANOVA) مناسباً؟

9- ما هي وظيفة مجلس المراجعة المؤسساتي؟

10- كيف تتأثر بعض النتائج البحثية بمعرفة الأفراد بهدف البحث؟ وكيف يمكن تلبية شرط الموافقة المطّلعة في هذه الظروف؟

11- ما هي تدبيرات الحذر المتوخى لضمان سرية الإجابات وخصوصية الأفراد في مشروع البحث ؟

12- أكتب بعض الأخطاء الشائعة التي ينبغي تجنبها في إعداد مقترح البحث.

الأجوبة ANSWERS

1- المقدمة، المنهجية / الطريقة، تحليل البيانات، أهمية الدراسة، الجدول الزمني والميزانية.

2- إذا لم تنفذ المقدمة بشكل جيد فسوف لن يميل القارئ إلى قراءة بقية المقترح.

3- في الفقرة الأولى من المقدمة.

4- أ. إن طلبة الصف الثالث ممن تعلموا القراءة في الصف الأول والثاني بالطريقة الكلية للغة سيحرزون معدلاً أعلى في اختبار كاليفورنيا للقراءة من طلبة الصف الثالث ممن تعلموا القراءة باستخدام أسلوب القراءة الأساسية.

ب. الطلبة الجدد وطلبة التخرج ممن أحرزوا درجات فوق الربيع الأعلى في اختبار xyz للاستعداد الميكانيكي قبل أن يصبحوا ممتهنين سوف يصنفون أعلى بكثير من قبل المشرفين بعد سنة من برنامج مهني ميكانيكي، من الطلبة الذين أحرزوا درجات أدنى من الربيع الأعلى.

ج. الدرجات في المجموعة الجزئية للرياضيات في اختبار التحصيل (SRA) لطلبة الصف السابع الذين لديهم معاملات ذكاء فوق (115) في (WISC) والذين تم تدريسهم بأسلوب الرياضيات الحديثة لمدة سنة، سوف تزيد على درجات الطلبة المماثلين لهم ممن تم تدريسهم بالأسلوب التقليدي.

5- متوسط مستويات القدرة المختلفة، وبيئات الصف المادية، والمدرسون، وأنواع ومقدار المواد هي بعض هذه العوامل. ويمكن للمرء ضبط بعض هذه المتغيرات عن طريق إجراء الاختبارات القبلية لمستوى القدرة، واختيار الصفوف ذات البيئات المتماثلة، وتدريب المدرسين إلى مستويات معينة من الكفاءة. واستخدام أعداد كبيرة من الصفوف المخصصة عشوائيا لمثل هذه الظروف سيقدم نوعا مختلفا من الضبط (control).

6- معامل الارتباط الثنائي أو الثنائي الأصيل أو الأصل أو معامل فاي (phi).

7- الوسط الحسابي.

8- المتغيرات المستقلة الاسمية والمتغيرات الفترية التابعة.

9- يحدد مجلس المراجعة المؤسساتي ما إذا كان الأفراد في تجربة مقترحة في خطر تحت طائلة الأنظمة الفيدرالية والمؤسسية، ويحدد ما إذا كانت الفوائد تفوق المخاطر، ويمنح أو لا يمنح موافقته على إجراء البحث.

10- المعرفة باستخدام معالجة غير اعتيادية أو تجريبية قد يؤثر على الحالة النفسية للأفراد / أو

توقعاتهم، بما قد يقلل من، أو يؤثر على، تأثيرات / نتائج المعالجة الفعلية. وفي مثل هذه الظرف، ينبغي القول للأفراد بأنهم سيطلعون على هدف الدراسة عند إكمالها.

11- لا يتعين على الأفراد تعريف أنفسهم إلا عند الضرورة ولا ينبغي تحديدهم كأشخاص في التقرير العام للدراسة إلا بعد أن يعطوا موافقتهم.

12- المشكلة تافهة.

المشكلة غير محدودة / محددة.

الصياغة الواسعة / العريضة جداً لأي من الأهداف أو الفرضيات أو الأسئلة.

الإجراءات تفتقر إلى التفاصيل.

اقتراح تصميم بسيط لمشكلة معقدة.

عدم أخذ المتغيرات ذات العلاقة بنظر الاعتبار.

REFERENCES المصادر

American Educational Research Association. (1991). Proposed ethical standards for AERA. *Educational Researchel:;* 20(9), 31-34.

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist,* 47; 1597-1611.

Code of federal regulations 45. (1993). (Part 34, pp. 293-294). Washington, DC: U.S. Government Printing Office.

Good, C.V., and Skates, D.E. (1954). *Methods ofresearch* (pp. 82-83). New York: Appleton.

Kromrey, J.D. (1993). Ethics and data analysis. *Educational Researchel:;* 22(4),24-27.

Smith, G.R. (1963). A critique of proposals submitted to the cooperative re- search program. In J.A. Culbertson and S.P. Hencly (Eds.), *Educational research: New perspectives* (chap. 17). Danville, IL: Interstate Publishers and Printers.



تحليل وتفسير وكتابة النتائج

Analyzing, Interpreting, And Reporting Result

أهداف تعليمية INSTRUCTIONAL OBJECTIVES

بعد دراسة هذا الفصل سيكون بوسع الطالب أن:

1- يصف دور الحاسوب في البحوث.

2- يدرج المبادئ التي ينبغي أن لا تغيب عن بال الباحث في ذهنه عند تفسير النتائج المتوقعة

3- يميز بين الدلالة الإحصائية والأهمية العملية.

4- يدرج المبادئ التي ينبغي على الباحث أن لا ينساها عند تفسير النتائج غير المتوقعة.

5- يوضح الظروف المتنوعة التي قد تؤدي إلى الاحتفاظ بالفرضية الصفرية.

6- يدرج بعض مبادئ التفسير التي ينبغي على الباحث أن لا ينساها عند الاحتفاظ بالفرضية الصفرية.

7- يكتب العناوين المناسبة لأطروحة أو أي تقرير بحثي آخر.

8- يصف طبيعة المضمون المراد إدخاله في كل جزء من رسالة أو أطروحة.

9- يبين الإجراء الذي يتبع في إعداد مقالة بحثية لمجلة.

10- يبين الإجراء المتبع في إعداد ورقة لقراءتها في مؤتمر للمختصين.

وحالما يتم جمع بيانات البحث، يعالج الباحث هذه البيانات بعدئذ، ويحلل النتائج ويفسر بدقة الاستنتاجات، وأخيراً يقوم بكتابة التقرير. إن مناقشة موجزة لكل من هذه المهمات معروضة في هذا الفصل.

معالجة البيانات PROCESSING THE DATA

الخطوة الأولى في معالجة البيانات المجمعة تكون بالعودة إلى المقترح بغية التأكد من الخطط الأصلية لعرض البيانات وإجراء التحليل الإحصائي. لقد تم مسبقاً اتخاذ القرار بشان الأسئلة التي سيتم الجواب عليها، وكيفية تعريف المتغيرات في هذه الأسئلة إجرائياً، وتحديد التصميمات البحثية والإجراءات الإحصائية التي سيتم استخدامها. أما كل ما يتبقى الآن فهو تنظيم البيانات الخام، وإجراء الحسابات الضرورية، وتفسير النتائج.

وغالباً ما تتسم المشاريع البحثية، في الوقت الحاضر، في التربية والعلوم السلوكية الأخرى بالتعقيد والأعمال الحسابية المتضمنة. ولتوفير الوقت والجهد فإن معظم الباحثين يستفيدون من التسهيلات الإلكترونية في الحسابات لتحليل بياناتهم. لذا من المناسب هنا أن نعرض مناقشة موجزة للحاسوب واستخدامه في معالجة البيانات.

استخدام الحواسيب في معالجة البيانات

The Use of Computers in The Processing of Data

مع ظاهرة انتشار وازدياد قوة الحواسيب الشخصية، إضافة إلى توفر البرمجيات الإحصائية لها، فربما يكون من المناسب القول الموثوق بأن أغلب الباحثين يقومون اليوم بتحليلاتهم باستخدام الحاسوب الشخصي. فللحواسيب أثر كبير على أنواع البحوث المنجزة. والبحث ذو العدد الكبير من المتغيرات والحسابات المعقدة، الذي كان نادراً، أصبح الآن مألوفاً.

ويجب أن يكون الحاسوب مبرمجاً بمجموعة من التعليمات لأداء أي عمل. وبينما كان الباحثون في الأيام الأولى من استخدام الحاسوب، يقومون ببرمجته بأنفسهم، فإن البرمجيات الإحصائية اليوم متوفرة بيسر للشراء أو للاستخدام من خلال إجازة الموقع. فعندما ينفذ البرنامج، فكل ما يحتاجه المرء فقط هو إدخال البيانات وتحديد المعْلمات / البارامترات للتحليل المرغوب - فالحاسوب سيؤدي التحليل المطلوب. إن مثل هذه البرمجيات تتراوح في حجمها من برمجيات ذات ثمن باهظ إلى برامج رخيصة متوفرة للطلبة، وتتراوح من البرامج معتدلة الحجم التي تلبي حاجات أغلب الباحثين إلى البرامج الكبيرة جداً التي يتم تطويرها للحواسيب الكبيرة وجرى تكييفها لاستخدام الحواسيب الشخصية. وبعض أمثلة البرامج الصغيرة هو برنامج MYSTAT وبرنامج الطالب SPSS و SYSTAT، SPSS-PC، و SAS فلها ما يقابلها في

الحاسوب الشخصي. أما البرامج الإحصائية الأخرى مثل SYSTAT و STATGRAPH فإنها قد كتبت خصيصاً للحاسوب الشخصي. وهناك معلومات أخرى عن هذه البرمجيات يمكن إيجادها في (Bowker's Software Encyclopedia, 1993).

وتسمح البرمجيات الإحصائية المنفذة في الحواسيب الشخصية للمستخدمين بتهيئة وإدارة تحليلاتهم بسهولة، والحصول على نتائج فورية، ثم إعادة أو تعديل التحليل عند الضرورة، وضبط نتاج التحليل بحيث يمكن استخدامه بشكل مباشر في تقرير بحثهم أو في الحصول على نتاج يمكن قراءته مباشرة في برامج المعالجة اللغوية المستخدمة في إعداد التقرير النهائي. إضافة إلى ذلك، فإن العديد من البرامج الإحصائية لها الآن قدرات متطورة لرسم المخططات والجداول والتمثيلات البيانية المعقدة بما يعطي للباحث خيارات متنوعة يستخدمها في عرض نتائج التحليل.

وعندما يؤخذ كل ذلك بنظر الاعتبار، فإن مرحلة تحليل بيانات البحث، التي كانت في الماضي واحدة من اصعب المراحل البحثية والأقل إمتاعاً في عمل البحث، لم تعد ذلك العمل الممل الذي كان. ففي الوقت الحالي، ومن خلال استخدام الحواسيب الحديثة فإنها لم تعد أسهل بل كذلك أكثر متعة.

التوصيات Recommendations

لابد من كلمة تحذير لازمة هنا. فالنتائج المتوفرة تكون جيدة بقدر جودة المعلومات التي جرى تغذية الحاسوب بها. فالحاسوب لا يرتكب الأخطاء الحسابية، إلا أنه إذا أدخلت أخطاء في البيانات أو أخطاء في المنطق إلى الحاسوب، فإن الحاسوب سيقوم بشكل أعمى ولكن بكفاءة بتحويل "نفاية" إلى نتائج. وهناك نوعان من الأخطاء في البرمجة. أخطاء لغوية مثل الأخطاء الهجائية للتعليمات، التي سيكتشفها الحاسوب حيث لن يقبل العمل في مثل هذه الحالة. أما الأخطاء المنطقية، مثل التعليمات غير الصحيحة، فإن الحاسوب سيقبلها بما قد يؤدي إلى نتاج باهظ لنتائج تخلو من المعنى.

أخيراً نود أن نطرح هذه التوصيات فيما يخص استخدام الحاسوب:

1- استفد بالكامل من الحواسيب في معالجة البيانات، ولكن قم بتقييم نقدي لناتج الحاسوب. فرغم أن الحواسيب لا ترتكب الأخطاء، إلا أنها أيضاً لا تصحح الأخطاء المرتكبة في إعداد البيانات للمعالجة أو البرمجة.

2- لا تستخدم أي برنامج بشكل أعمى مما تجده في كتيبات البرامج. فيجب أن تقرر أي نوع من الإجراءات الإحصائية يكون أكثر ما يناسب لجملة معنية من البيانات.

مبادئ التفسير
PRINCIPLES OF INTERPRETATION

حالما يتم جمع البيانات البحثية ويجرى التحليل الإحصائي يمكن للباحث أن يمضي إلى مهمة التحدي لتفسير النتائج. فالإضافة إلى المعرفة هي النقطة الرئيسية لمشروع البحث. وعند التوصل إلى مرحلة التفسير، يمكن للباحث أن يبين ما تعلمه في المشروع وكيف أن هذه المعرفة تلائم مجموعة المعارف في الحقل.

دور المقترح في تسهيل التفسير
The Role of The Proposal in Facilitating Interpretation

يجب إقامة الأساس المناسب لتفسير نتائج دراسة، بشكل منتظم خلال كل مرحلة من تطور المقترح، حتى قبل بدء البحث الفعلي. وبعد أن نضع في الذهن خلال الدراسة مكونات البيانات وما قد تفصح عنه، يعد الباحثون أنفسهم لتفسير بياناتهم وجعلها ملائمة لمجموعة المعارف.

إن الخطة المعدة جيدا والمعبر عنها بالمقترح الكامل والعميق، يمكن لها أن تعطي نتائج يمكن تفسيرها بسهولة ودلالة متميزة. وإذا تم ترتيب الدراسة بطريقة يتم فيها التعبير عن نتائج الفرضيات بملاحظات موثوقة، فلا بد أن يكون تفسير الملاحظات وقيمتها واضحين.

أهمية الالتزام بالخطة الأصلية
The Importance on Keeping to the Original Plan

في بحوث اختبار – الفرضية، ولدى القبول بالمقترح وبدء انطلاق المشروع، يجب تنفيذ الدراسة كما هو مخطط لها بالضبط. ولهذه القاعدة مضامين أخلاقية وعملية.

ولتوضيح المضامين الأخلاقية افترض أن السيد "س" وهو مدرس لغة أجنبية، قد طور نظاما بعد أن انفق المال والجهد لتدريس الفرنسية وهو يعتقد أنه افضل من الطرق الموجودة. ولاختبار فعالية هذه الطريقة، فإنه يكوّن مجموعة تجريبية يتم تدريسها بطريقته ومجموعة ضابطة يتم تدريسها بطريقة أخرى. ويقوم بابتكار سلسلة اختبارات أسبوعية لتحصيل الفرنسية يتم استخدامها كمتغير تابع. وافترض أنه يكتشف في الأسابيع القليلة الأولى أن وسطي درجات الاختبار للمجموعتين متطابقان تقريباً. ونظراً لاستثماره الكبير في طريقته الخاصة، فإنه يجد من الصعوبة التصديق أنها ليست افضل من الطريقة الأخرى، لذا يقرر مراقبة الصفين ليرى "ما وقع من خطأ". ويكتشف أن المجموعة التجريبية تبدي معرفة وتقديراً أكبر للحياة والثقافة الفرنسية. ونظراً لأنه مصمم على إيجاد فرق بين الدرجات فإنه يقرر تغيير متغيره التابع إلى درجات في الاختبارات حول الحياة والثقافة الفرنسية.

إن مثل هذا التغيير غير أخلاقي. فإذا أخذنا مجموعتين عشوائيتين، فبوسع المرء أن يجد دوماً من خلال الصدفة وحدها تفوقاً في أي من المجموعتين إذا تأمل بشكل مطول بما يكفي. فإذا لم تكن المجموعة التجريبية متفوقة في الحياة والثقافة الفرنسية فإنها قد تبدو متفوقة في الطلاقة اللفظية، أو مهارات الإنصات، أو السلوك أثناء الواجب، أو بأي متغير آخر. فيجب على مدرس اللغة أن ينفذ التجربة كما هو مخطط لها وأن لا يغير المتغير التابع. وينبغي على الباحث ذكر الأدلة التي تشير إلى العلاقة بين الطريقة وتقدير الحياة والثقافة الفرنسية، غير أنه يجب التوضيح لدى القراء أن تلك لم تكن علاقة مفترضة، وعليه قد تكون بفعل الصدفة. إنه ليس أخلاقياً استبعاد المتغيرات المستقلة والتابعة التي لا يبدو أنها "تعمل بنجاح" أو إضافة متغيرات جديدة واعدة. فيجب ترك مثل هذه التغيرات لدراسات مستقبلية.

وبالطبع، فإن هذا التحذير ينطبق على البحث الذي يجرى فيه اختبار فرضيات مقررة مسبقا. أما في البحث النوعي، فإن الفرضيات تبرز مع تقدم الدراسة ويعاد تحديدها عندما يكون الأمر مناسباً. وكذلك في البحث الوصفي – كالدراسات التطورية النمائية وتحليل الاتجاهات / التيارات التي لا تبدأ بفرضية – فليس هناك من سبب يدعو الباحث لعدم إضافة متغيرات.

إن إضافة متغيرات جديدة هو أمر غير حكيم من وجهة نظر عملية. فمثل هذا الأسلوب قد يربك نتائج دراسة اختبار – الفرضية، ويشوش مضمون النتائج. فالباحثون غالبا ما يقعون تحت إغراء إضافة متغيرات جديدة مثيرة تبرز على نحو غير متوقع في دراستهم. ومع ذلك، فإن القاعدة النظرية لتفسير هذه المتغيرات لم توضع بعد، وثانية فإن أفضل نصيحة للباحثين تكون بتركها لدراسات أخرى.

تفسير النتائج المتوقعة

Interpretation of Expected Results

من المفهوم أن الباحثين يكونون مسرورين عندما تكون نتائج الدراسة مناسبة للإطار الذي تم بناؤه مسبقاً وأن بوسع التفسير أن يمضي كما هو متوقع. لقد "نجحت" الدراسة وثمة اتفاق بين الأساس المنطقي والنتائج.

وهناك فقط بضع كلمات تحذير نحتاج إلى تطبيقها في مثل هذه الحالة:

1- لا تجعل تفسيراتك تتجاوز حدود المعلومات. وقد يبدو ذلك أمراً واضحاً بصورة جلية، إلا أن الباحثين غالبا ما يستثارون عندما تأتي النتائج كما هو متوقع لها، فيستمدون استنتاجات لا أساس لها من الصحة في البيانات. وحتى في البحوث المنشورة، يجد المرء أحياناً تفسيرات أكثر مما تسوّغه البيانات.

2- لا تنس حدود الدراسة. وبالطبع، فإنه ينبغي تعيين حدود الدراسة مسبقاً – وهي

حدود متأصلة في ثبات وصدق الأدوات المفتقرة للكمال، وهي حدود تنشأ عن التقييدات في المعاينة ومشكلات الصدق الداخلي، وما إلى ذلك.

3- تتطلب الأخلاقيات أن يذكر الباحث مشكلات الصدق الداخلي التي قد تؤثر في النتائج. فإذا كانت المتغيرات غير التجريبية، رغم أفضل جهود الباحث، بشكل خاص غير مضرة بالمجموعة التجريبية بينما كانت ضارة بشكل خاص في المجموعة الضابطة، فيجب ذكر هذه الشروط وأخذها بنظر الاعتبار في تفسير النتائج. (مثلاً، رغم التخصيص العشوائي للمدرسين في مجموعات، فقد تشمل المجموعة التجريبية على أكبر المدرسين خبرة وتشتمل المجموعة الضابطة على أقل المدرسين خبرة).

4- تذكر أن الدلالة الإحصائية تعني، فقط، انه لدرجات مناسبة من الحرية لا يحتمل أن تكون النتائج من فعل الصدفة. فالدلالة الإحصائية لا تعني أن النتائج مهمة حسب المعنى المقبول عموما للكلمة – أي، مهم، ذو معنى، أو هام جداً. لا تفترض أن الدلالة الإحصائية تضمن شأناً متميزاً لاستنتاجاتك.

افترض أن مجموعتين متكافئتين قد تعرضتا لنظامين مختلفين في تعلم التهجئة على فترة سنتين. فأولئك الذين يستخدمون النظام "أ" يحصلون على زيادة في النمو بمتوسط 2.15 سنة في الاختبارات المقننة خلال التجربة، بينما يزيد الذين يستخدمون النظام "ب" بمتوسط حاصل 2.20 سنة في ذات الفترة. فإذا كانت المجموعات كبيرة و / أو كانت الفروقات داخل المجموعات صغيرة، فإن الفروقات بين الأوساط ستكون ذات دلالة إحصائية. غير أن فرق نصف الشهر على فترة سنتين هو فرق دون معنى نسبيا بالإطار العملي. فإذا كان النظام "ب" أكثر كلفة في إطار وقت الطالب، أو وقت المدرس، أو المواد فإن المدرسين سيكونون غير حكيمين في تبنيه لأنه ببساطة أعطى كسباً ذا دلالة إحصائية اكبر من النظام "أ". ولو كان النظام "ب"، من ناحية أخرى أقل كلفة، فسوف يميل المدرسون إلى تفضيله لأن نتائجه مماثلة لنتائج النظام أ حسب الإطار العملي.

ويجب تحديد الأهمية المحتملة أو المعنى المتميز المحتمل للنتائج في المقترح قبل بدء الدراسة. فالدراسة لا تكون مهمة إن لم توفر المعلومات ذات المعنى التي ستضاف إلى مجموعة المعارف القائمة، أيا تكون الدلالة الإحصائية للنتائج.

تفسير النتائج السلبية

Interpretation of Negative Results

الباحثون الذين يجدون النتائج مناقضة لما افترضوه، فإنهم غالباً، يكتشفون أمراً مفاجئا يتعلق بمواطن الضعف في دراستهم. فتفسيرهم للنتائج يكون بمثل الاعتراف. فالأدوات كانت غير وافية لقياس المتغيرات المعنية، كما كانت العينة صغيرة جدا ولذا كانت غير ممثلة بحيث أن

النتائج لا يمكن تعميمها بشكل صادق على مجتمع مستهدف متميز، وهكذا. ويكشف الإدراك المتأخر عن مشكلات الصدق الداخلي بما يوضح سبب عدم وصول الدراسة إلى "ما يجب".

وبالطبع، فإن أياً من هذه الأمور أو كلها قد تكون صحيحة، فينبغي ذكر نقاط ضعف أية دراسة مهما كانت النتائج. وعلى أية حال، فإن البحث يعتبر دائما مغامرة في المجهول، لذلك لا يوجد "ما يجب أن يكون" كشيء مطلق. فالباحث يتكهن بنتائج الدراسة المتوقعة على أساس النظرية، والاستقراء، والتجارب، ونتائج البحوث السابقة. وإذا كانت هذه مقنعة بعدم وجود شك على الإطلاق بشان نتائج الدراسة هذه، فتكون الدراسة دون معنى / تافهة في المقام الأول.

وعندما تتعهد القيام بدراسة، نوضح ضمنا أن النتيجة قضية حدس وليست يقينا. وعندما نكمل مقترحاً، فنحن، من الواضح، نصرح أننا سوف نبحث دون تحيز عن تحديد الحالة الحقيقية للأمور مع افضل الأدوات والإجراءات المتيسرة للاستخدام من أجل ذلك الهدف. وعليه. نحن ملزمون بقبول وتفسير بياناتنا مهما كانت هذه البيانات. وعندما تتعارض النتائج مع الأساس المنطقي النظري للدراسة، فإن الجزء الخاص بالمناقشة في تقريرنا، ينبغي له أن يتضمن إعادة النظر في النظرية الأصلية في ضوء الاستنتاجات. وغالباً ما يكون الباحثون غير راغبين في عرض تفسير البيانات التي تتعارض مع بحوث سابقة أو مع نظرية راسخة. ومع ذلك، فقد تكون نتائجهم صحيحة والنتائج السابقة خاطئة. إن تقدم علم التربية سوف يُعوّق إذا لم يرغب الباحثون في نشر نتائجهم التي لا تتفق مع النتائج المذكورة في الدراسات السابقة. إن النتائج المتناقضة توضح أن مسالة لم تحسم وقد تثير بحوثا أخرى. فالبحث الإضافي أو صياغة نظرية قد يوفق بين النتائج المتناقضة. فالنظرية مؤقتة / غير نهائية وينبغي ألا يعوق الباحثين من إعطاء تفسير صريح لما تم إيجاده.

إن إعادة النظر في الأساس النظري لدراسة معينة يعود إلى الجزء الخاص بالنقاش. فيتعين على المرء أن لا يرجع إلى الخلف ويعيد كتابة الأدبيات ذات الصلة والفرضيات في الأجزاء الخاصة في التقرير.

تفسير النتائج عند الاحتفاظ بالفرضية الصفرية

Interpretation of Results When The Null Hypothesis Is Retained

طالما أن بالإمكان الاحتفاظ بالفرضية الصفرية لأسباب مختلفة، فإن تفسير مثل هذه النتيجة قد يكون صعبا بشكل خاص. وقد تحدث الفرضية الصفرية لأن: (1) الفرضية الصفرية صحيحة في الواقع. فقد لا تكون هناك علاقة بين المتغيرات. وليست المعالجة التجريبية أكثر فاعلية من المعالجة الضابطة. (2) الفرضية الصفرية خاطئة، غير أن مشكلات الصدق الداخلي قد شوشت الاستقصاء بشكل رديء إلى درجة تعذر ملاحظة العلاقة الفعلية بين المتغيرات. (3) الفرضية الصفرية خاطئة إلا أن تصميم البحث يفتقر "للقوة" في رفضها.

إن أياً من هذه الحالات قد تكون هي المعنية، إلا أن الباحث لا يعرف الصحيح منها وعليه لا ينبغي له أن يدعي بأحقية أية واحدة منها لتفسير للنتائج.

ومن غير الصحيح عرض فرضية صفرية ثم الاحتفاظ بها كدليل على عدم وجود علاقة بين المتغيرات. وينعين تفسير فرضية صفرية صار الاحتفاظ بها بكونه يعود إلى افتقار للدليل على صحة أو خطأ الفرضية. إن إعلاناً تجارياً عن معجون الأسنان شائع الاستخدام ذكر أن الاختبارات تبين أن معجون أسنان خاص لم يتفوق عليه معجون آخر في تقليل تلف الأسنان. وبتفسير مصطلح "لم يتفوق عليه" بمعنى "عدم وجود دلالة إحصائية" يمكن أن نتصور اختباراً استخدم فيه عدد صغير من الأفراد و / أو احتوى على مشكلات صدق داخلي متعددة. فإذا كانت الفرضية الصفرية هي النتيجة المرغوبة للتجربة فإن من السهل الترتيب لمثل هذه النتيجة.

هناك خطر من أن يقع الباحثون الذين يصبحون متيّمين كثيرا بفرضيتهم التجريبية تحت إغراء استبعاد فرضية صفرية محتفظ بها. إنهم يذكرون مشكلات الصدق الداخلي ويعلنون بأن النتائج ستكون دالة بالتأكيد لو أن تلك المشكلات فقط غير المتوقعة لم تفسد التجربة. وبالطبع، ينبغي على المرء أن يذكر جميع مشكلات الصدق الداخلي التي تبرز في الدراسة، غير أنه يتعين على المرء عدم استخدامها في تفسير عدم مسئوليتهم عن النتائج المحبطة. وقد يقترح المرء بحثاً إضافياً مخططاً بطريقة تتجنب مواجهة مشكلات الصدق الداخلي، إلا أنه لا يزال على المرء وجوب بيان الاحتفاظ بفرضية صفرية كافتقار للدليل ولا أكثر.

تذكر من الفصل 6 بأن "قوة" التجربة تشير إلى القدرة الإحصائية لرفض فرضية صفرية حين تكون، في الحقيقة، خاطئة. وهذه القوة تعتمد على (دالة في) حجم العينة، وتجانس الأفراد بما يتعلق بالمتغير التابع، وثبات أدوات القياس المستخدمة، وطبيعة الإجراء الإحصائي المستخدم لاختبار الفرضية وكذلك حجم التأثير. وينبغي على الباحثين أخذ هذه العوامل بنظر الاعتبار عند التخطيط لتجربة. فيجب أخذ قوة التجربة بنظر الاعتبار لدى التخطيط لتجربة. فيجب اعتبار قوة التجربة في التخطيط للدراسة. ولا يجب إدخالها في نهاية الدراسة لتفسير عدم مسئوليتهم عن الافتقار إلى الدلالة الإحصائية. مثلا، لا ينبغي على المرء القول أن "النتائج ستكون ذات دلالة إحصائية لو كانت العينة أكبر".

مع استثناءات نادرة، فإن التفسير المشروع الوحيد للاحتفاظ بفرضية صفرية هو "الدليل الكافي لنتيجة لم تتم ملاحظته". وبالطبع، إن كان المرء يدرس مجتمعاً إحصائياً صغيراً وكان بوسعه إجراء إحصاء تام لذلك المجتمع، فإن الاحتفاظ بفرضية صفرية يمكن تفسيره بشكل مشروع كافتقار للعلاقة بين المتغيرات ضمن ذلك المجتمع الخاص. ويتطلب الاحتفاظ بفرضية صفرية مصداقية إن أمكن التوضيح بأن الدراسة خالية من مشكلات الصدق الداخلي وأن لها قوة كافية لرفض الفرضية الصفرية بحجم تأثير تافه نسبيا.

تفسير علاقات غير مفترضة
Interpretation of Unhypothesized Relationships

لقد أكدنا سابقا على أنه لا ينبغي على الباحث أن يستبعد فرضية خلال إجراء دراسة معينة بغية اتباع سبل واعدة اكثر تطرح نفسها خلال مسار الدراسة. ولا يعني ذلك وجوب تجاهل أية علاقات غير مفترضة قد تجرى ملاحظتها في تنفيذ الدراسة. وعلى العكس، ينبغي تسجيلها وتحليلها بذات النشاط المستخدم في متابعة العلاقات المفترضة. وخلال تاريخ العلوم، برهنت الكثير من اكتشافات المصادفة على أهميتها.

وعلى أية حال، يجب النظر إلى مثل هذه الاستنتاجات دائماً بشك أكثر من الاستنتاجات المرتبطة بشكل مباشر بالفرضية، لأن هناك إمكانية كبيرة نسبيا بوجود علاقة غير مفترضة زائفة ستظهر في دراسة ما. وينبغي ذكر مثل هذه العلاقات، غير أنه ينبغي اعتبارها طارئة على قوة الضغط الرئيسي للبحث. وينبغي جعلها موضوع دراسة مصممة بشكل خاص لاستقصائها قبل إمكانية استخدامها كأساس لاستنتاجات معينة.

THE RESEARCH REPORT تقرير البحث

إن نتائج المشروع البحثي هي ذات قيمة قليلة ما لم يمكن إيصالها إلى الآخرين. وعليه، فإن معرفة الإجراءات التي تنطوي عليها كتابة تقرير البحث مهمة لجميع الباحثين. وهدف هذا الجزء هو تقديم دليل عام لتنظيم وعرض التقرير. ومن أجل القواعد الخاصة بالأسلوب والشكل العام يجب الرجوع إلى كتيب في الأسلوب (العديد منها مدرج في نهاية الفصل).

ويقوم الباحث، في تقرير البحث، بإيصال إجراءات واستنتاجات البحث كما يناقش أيضاً مضامين الاستنتاجات وعلاقتها بالمعرفة الأخرى في الحقل.

ونظراً لأن مختصين منهمكين بأعمالهم سيقرأون هذا التقرير، فلا بد أن يكون موجزاً ومنظماً بشكل منطقي قدر الإمكان. فلا مكان للقصص، وروايات التجارب الشخصية، والمحادثات الجدلية في تقرير البحث. ولا يعني ذلك ضرورة أن يكون التقرير مملاً ومتحذلقاً. فإذا قام الباحث بإجراء الدراسة بروح حماسية فإن هذه الروح تميل إلى الانتقال ما بين السطور.

وبما أن هدف التقرير هو عرض البحث وليس شخصية المؤلف، فالأسلوب لا ينبغي له أن يكون شخصيا. ولمواكبة ذلك فإن ضمائر الشخص الأول لا تستخدم مطلقا. وهكذا لا يتعين على أحد أن يكتب "أنا اخصص عشوائيا أفراد المجموعتين في المعالجة". بل عليه أن يقول "جرى تخصيص الأفراد عشوائياً للمجموعتين" ورغم الحماسة الطبيعية حول أهمية العمل، لا ينبغي على المؤلف أن يتباهى بذلك بل يجب أن يترك التقييم إلى القراء وذريتهم.

لقد تطورت طريقة رسمية وموحدة لعرض التقارير البحثية. فرغم أن هذه الشكليات قد تبدو مثبطة حسب النظرة الأولى، إلا إن الممارسة تشير إلى أنها تخدم هدفاً مفيداً. فمن المهم ترتيب التقارير البحثية بطريقة يعرف القراء بالضبط أين يجدون تلك الأجزاء الخاصة التي يبحثون عنها. وإلا فإن الوقت سيتبدد بحثا عن المعلومات المناسبة. إضافة إلى ذلك، فإن وجود شكل عام راسخ يستبعد الحاجة لابتكار صيغة خاصة بالمرء. وخلال مناقشة هذا الموضوع، فسوف يتضح أن الشكل العام القائم سوف يتبع بصورة منطقية الخطوات في المشروع البحثي المعروضة في الفصول الأولى.

وقد يعرض التقرير البحثي كـ: (1) رسالة أو أطروحة، (2) مقالة بمجلة، (3) ورقة مؤتمر. وثمة أسلوب مختلف مطلوب في كل من هذه الحالات.

الرسالة أو الأطروحة The Thesis Or Dissertation

إن لدى العديد من الجامعات كتيباً مفضلاً يصف بالتفصيل الشكل الذي تطلبه الجامعة. وبالنسبة للطلبة الذين لديهم حرية الاختيار، فإن العديد من كتيبات الأسلوب مدرجة في نهاية هذا الفصل. وعندما يتم اختيار الكتيب، فإنه يجب كتابة التقرير كله حسب هذا الأسلوب وفقا لتوصياته.

يدرج المخطط التالي السياق والعناصر العامة الموصوفة في أغلب كتيبات الأسلوب:

1- صفحات تمهيدية
- أ. صفحة العنوان
- ب. صفحة الإجازة / القبول
- ج. الشكر أو التصدير
- د. قائمة المحتويات
- هـ. قائمة الجداول
- و. قائمة الأشكال

2- المتن(*) (Text)
- أ. المقدمة
  1. صياغة المشكلة والأساس المنطقي للدراسة
  2. الأهداف
  3. تعريف المصطلحات

(*) المتن (text): يشير إلى الجزء الأساسي (النص الفعلي) الذي يعرضه الباحث مجرداً من الهوامش والصفحات التمهيدية والصفحات الملحقة – (المراجع).

4. الأدبيات ذات الصلة

ب. الطرق والنتائج
1. الأفراد
2. الإجراءات
3. الأدوات
4. عرض وتحليل البيانات

ج. مناقشة النتائج (results)
1. تفسير الاستنتاجات
2. المضامين
3. التطبيقات

د. الاستنتاجات والخلاصة
1. الاستنتاجات ، النتائج النهائية (conclusions)
2. الخلاصة (summary) / موجز الدراسة

3- صفحات ملحقة
أ. ثبت المراجع
ب. الملاحق
ج. السيرة (عند الحاجة)
د. الخلاصة (Abstract) / موجز قصير جداً

الصفحات التمهيدية Preliminary Pages

إن إعداد الصفحات التمهيدية هي مسألة تتبع بشكل كبير قواعد كتيب الأسلوب. ومع ذلك فأن أحد جوانب هذه الصفحات الذي يحتاج إلى إيضاح إضافي عند هذه النقطة هو عنوان الدراسة نفسها.

فالعنوان، ينبغي أن يصف بشكل مختصر ، بقدر الإمكان، الطبيعة المحددة للدراسة. مثلاً، خذ (أ) دراسة الأطفال المحرومين ثقافياً التي تقارن استعداد القراءة لمن شاركوا في برنامج مشروع ميزة البداية(*) مع استعداد مجموعة مناظرة من الأطفال الذين ليست لديهم خبرة رسمية في ما قبل المدرسة، و (ب) العنوان "مقارنة درجات اختبار الاستعداد للقراءة للأطفال المحرومين

(*) مشروع ميزة البداية (Head-Start Project): مصطلح يشير إلى مشروع/ برنامج تربوي لمساعدة الطلبة المحرومين ثقافياً (أو غيرهم من فئات خاصة) من تحسين مستواهم/ استعدادهم قبل دخول المدرسة، ويشير بشكل عام إلى إعطاء ميزة/ أفضلية لأفراد على آخرين من أقرانهم - (المراجع).

ممن انضموا إلى صفوف ميزة البداية لستة أسابيع أو أكثر مع أطفال مماثلين ليست لديهم خبرة ما قبل المدرسة". فبينما ينقل هذا العنوان عما تدور عليه الدراسة، إلا أنه طويل جدا. فعبارات مثل "مقارنة"، "دراسة حول" و "بحث في" هي عبارات فائضة. علاوة على ذلك، فإن أغلب القراء المتوقعين سيعرفون أن "مشروع ميزة البداية" هو خبرة لما قبل المدرسة مصمم للأطفال المحرومين ثقافياً. ومع ذلك إذا تطرفنا وأعطينا عنواناً موجزاً أو غامضاً جداً لنقل طبيعة الدراسة لكان ذلك خطأ أكثر فداحة. فالعناوين الغامضة أو المختصرة تفرض على القارئ أن يبحث في المقال بغية تحديد ماهيته. فعناوين مثل "ميزة البداية والاستعداد" أو "القراءة بين المحرومين" توضح هذا العيب. "فالعنوان ينبغي أن يحدد المتغيرات الرئيسية والمجتمعات المعنية". فلا ينبغي إدخال التعريفات الإجرائية للمتغيرات الرئيسية ووصف العينات في العنوان.

ونظراً لأن العنوان الصحيح سيضمن الفهرسة الصحيحة، فإن استراتيجية مفيدة للباحثين تكون بتقرير، تحت أية كلمات رئيسية يريدون فهرسة دراساقم، والعمل من هناك نحو العنوان الموجز. ففي مثالنا، ستكون الكلمات الرئيسية المهمة للفهرسة هي "الاستعداد للقراءة" و "مشروع ميزة البداية". لذا، فالعنوان المناسب قد يكون "استعداد القراءة لأطفال مشروع ميزة البداية وأقرانهم ممن لم ينضموا للبرنامج". ويعتبر هذا العنوان موجزاً بشكل معقول وهو يعطي القارئ المأمول مؤشراً صحيحاً عما تكون عليه الدراسة.

أما العناوين العاطفية مثل "يتعين علينا توسيع مشروع البداية الأساسية" أو "لكي لا ندع المحرومين أن يصبحوا قراءاً ضعفاء" فيجب تجنبها بأي حال. فالقارئ المأمول لن يتوقع استنتاجات البحث تحت مثل هذه العناوين بل يتوقع مقالات نظرية تحاول إغراءه بوجهة نظر.

المتن Text

المقدمة: يشمل الجزء الخاص بالمقدمة على كل ما حدث في عرض أساس البحث. وهو يتكون عادة من مواد معدة مسبقا للمقترح مع بعض التغيرات الطفيفة نسبياً. فصياغة المشكلة وتسويغ الدراسة يظلان على حالهما تماماً مثل صياغة الأهداف، وتعريف المصطلحات، ومراجعة البحوث ذات الصلة. وعلى أية حال، فأن التقرير الموجز عن مصادر البيانات والطرق المستخدمة تكتب بصيغة الماضي، وليس المستقبل المستخدم في المقترح الأصلي. أما مراجعة البحث ذات الصلة فتعرض عادة كفصل منفصل في جزء المقدمة.

الطرق والنتائج: هناك أربعة أصناف من المعلومات مشمولة في هذا الجزء من التقرير:

1- الأفراد: يجب إعطاء وصف تفصيلي للعينة. وهذا يمكّن القارئ من الحكم على الصدق الخارجي المحتمل للمجتمع الإحصائي الخاص بالبحث. وينبغي تحديد/ تعريف المجتمع الذي استمدت منه العينة كما ينبغي تحديد حجم وطريقة المعاينة. أما نوع المعلومات المقدمة في وصف العينة فإنه سيختلف من دراسة إلى أخرى، إلا أنه يمكن للمرء بشكل عام أن يقرر أية معلومات سيضمنّها عن

طريق النظر في المتغيرات التي قد تؤثر على درجات المتغير التابع/ المعيار في الدراسة.

2- **الإجراءات**: يجب أن يكون ذكر الإجراءات كاملاً بما يكفي بحيث أنه إذا أراد أحد إعادة/ استنساخ الدراسة سيجد جميع المعلومات الضرورية هناك. إن إحدى سمات الطريقة العلمية هي إمكانية تأكيد الاستنتاجات عن طريق تكرار الإجراءات ومعلومات الرصد الضرورية للاستنساخ. أما تصميم الدراسة، وعدد المجموعات (إن كانت الدراسة تجريبية)، ومعالجة الأفراد، والمعلومات الأخرى ذات الصلة فكلها مشمولة في هذا الجزء.

3- **الأدوات**: ينبغي أن يحدد تقرير البحث جميع أدوات القياس وأنظمة الرصد/ الملاحظة المستخدمة في تنفيذ الدراسة، ويمكن أن تكون المواصفات موجزة عند استخدام المقاييس المقامة سابقاً، لأن المراجع ستحتوي على معلومات ذات صلة حول مثل هذه الأدوات. فإذا تم بناء أدوات خاصة بالدراسة فيجب تقديم وصف مفصل لهذه الأدوات مع دليل للثبات والصدق ومناقشة إجراءات وضع الدرجات. وإن كانت الدراسة مسحاً، فيجب ذكر النسبة المئوية للإعادة وأية جهود أخرى لتحديد سمات غير المستجيبين.

4- **عرض وتحليل البيانات**: إن أحد الأساليب الموصى بها للعرض والتحليل الإحصائي للبيانات هو تنظيم النقاش حول الفرضيات - أي إعادة عرض الفرضية الأولى وعرض الاستنتاجات الخاصة بها وإعادة هذا الإجراء لكل فرضية حسب الدور.

وقد تستخدم الجداول والأشكال بشكل مفيد لعرض البيانات بشكل أوضح وأكثر إيجازاً مما هو ممكن إذا عرضت المعلومات ذاتها في المتن. إن أغلب الكتيبات الخاصة بالأسلوب تقدم أمثلة عن الأنواع المستخدمة الشائعة للجداول والأشكال وتعليمات بنائها. فالجدول المعد بشكل جيد يمكن أن يعطي القارئ نظرة عامة موجزة عن البيانات.

أما الجداول التي يجري بناؤها لدى تنفيذ البحث، فإنه عموماً لا يمكن دمجها مباشرة في التقرير. فمثلا، قد يكون لدى المرء بعد إكمال دراسة معينة قائمة أبجدية بالأفراد في دراسة معينة ودرجاتهم حسب مقاييس معيار/ (متغير تابع). وبدلاً من عرض هذه القائمة كما هي، يمكن أن يبني المرء جدولاً بالمعلومات بصيغة مختصرة. (أما جداول البيانات الخام الأساسية فقد تضم في الملاحق إذا كان ثمة شعور بإنها قد تساهم في الفهم). إن الجدول الأول في التقرير يلخص عادة البيانات الوصفية مثل الأوساط الحسابية، والانحراف المعياري، والارتباطات، والنسب المئوية. وتعرض جداول أخرى لاحقاً نتائج تطبيق الإحصاء الاستدلالي واختبارات الدلالة للبيانات. مثلاً، قد يستخدم جدول مختصر لعرض نتائج تحليل التباين.

ومن المرغوب، ترتيب البيانات بطريقة توضح علاقة البيانات بفرضيات الدراسة. أما الطريقة الفاعلة فتتم بعرض البيانات في جداول وأشكال يصاحبها نص كاف للإشارة إلى الاستنتاجات الأكثر أهمية والنتائج المثيرة. ومن المهم بشكل خاص ربط المعلومات في الجداول بالفرضيات.

ويجب توضيح الأساس الإحصائي لتحليل النتائج بشكل جلي. ومن المناسب ضم المعالجة الإحصائية مع عرض البيانات.

مناقشة النتائج: تفسر النتائج مرة أخرى من خلال علاقتها بالفرضيات (أو الأسئلة) كما تناقش مضامين وتطبيقات الدراسة.

1- تفسير النتائج: ربما يكون الجزء الأصعب، والأكثر فائدة أيضاً، من التقرير هو تفسير الباحث للنتائج. وهذا الجزء يربط الاستنتاجات النظرية والبحوث السابقة في هذا المجال وكذلك بإجراءات البحث.

2- **المضامين:** إن مساهمة النتائج بالمعرفة في الحقل العام للدراسة قضية ينبغي مناقشتها في هذا الجزء. فالباحث يوضح هنا كيفية تعديل النتائج التي لها علاقة بالنظريات ويقترح دراسات أخرى تلي بشكل منطقي.

3- التطبيقات: إن توضيح تطبيق النتائج يساعد القراء على فهم كيفية تطبيق النتائج في الممارسة.

إن الأجزاء الخاصة بالمضامين وتطبيقات النتائج لا تتطور في الغالب بشكل كاف حيث يفترض أنها ستكون واضحة لدى القارئ كما هي واضحة لدى الباحثين. وفي الواقع، فإنه ربما حصل الباحثون أثناء إجراء الدراسة على فهم للمشكلة أعمق مما يفترض بأن أغلب القراء حازوا عليه. وعليه قد يتوقع المرء أن تكون التفسيرات أكثر معنى من التفسيرات التي يصنعها القراء لأنفسهم.

الاستنتاجات والخلاصة (Conclusions and Summary): تشكل أجزاء الخلاصة والاستنتاج قمة / أوج التقرير.

1- **الاستنتاجات:** ينبغي أن تتحدد مناقشة النتائج النهائية التي دلت عليها استنتاجات البحث، بتلك الجهات التي لها دعم مباشر في تمويل البحث. وغالباً ما يقع الباحثون تحت إغراء استنتاج الكثير جداً. إن الفرضيات توفر الإطار المناسب لصياغة الاستنتاجات، أي أنه يتعين على الباحثين الإشارة في هذا الجزء، إلى ما إذا كانت الاستنتاجات تدعم الفرضيات.

من المهم التمييز بين النتائج (results) والنتائج النهائية / الاستنتاجات (Conclusions).(*)

(*) النتائج (results) والاستنتاجات/ النتائج النهائية/ الخاتمة (conclusions): تستخدم كلمة (نتيجة) في اللغة العربية للمصطلحين (conclusion, result)، ولا مانع من ذلك إذا أمكن التمييز بينهما كما توضح الفقرة أعلاه، باعتبار الأولى (results) تشير إلى نتائج والأخرى (conclusion) تشير إلى استنتاج / استدلال/ نتيجة نهائية - (المراجع).

فالنتائج (results) هي ملاحظات مباشرة ملخصة ومدمجة عن طريق التحليل الإحصائي مثل مقارنة وسطي مجموعتين. أما الاستنتاج / النتيجة النهائية (conclusion) فهو استدلال (inference) مستند إلى النتائج، ويعبر عنه في إطار فرضية الدراسة، مثل معالجة مجموعة وكونها أكثر تأثيراً من معالجة المجموعة الأخرى. مثلاً، قد تنتهي دراسة معينة بملاحظة أن وسط درجات اختبار التهجئة لدى الطلبة الذين تم تدريسهم التهجئة حسب الطريقة (أ) هو أعلى من وسط الطلبة الذين تم تدريسهم حسب الطريقة (ب). إن النتيجة / الاستنتاج بأن الطريقة (أ) أكثر فاعلية من الطريقة (ب) هو ليس نتيجة (result) مباشرة للدراسة بل هو استدلال يستند إلى نتائج (results) الدراسة.

وقد يُدخل الباحثون مناقشة موجزا لأفكارهم حول مضامين استنتاجاتهم وتوصياتهم من أجل تطبيقات محتملة للاستنتاجات. وقد تشير كذلك إلى أية أسئلة جديدة للبحث نشأت عن الدراسة.

2- **الخلاصة (Summary)**: نظراً لأن الخلاصة ستقرأ بشكل واسع أكثر من بقية أجزاء التقرير، فيجب أن تكون صياغتها بشكل خاص واضحة وموجزة. وعادة ما تتضمن الخلاصة إعادة صياغة موجزة للمشكلة والمظاهر الرئيسية للطرق والاستنتاجات الأكثر أهمية. ولدى إكمال مسودة هذا الجزء، ينبغي على المؤلف أن يدققها بعناية ليحدد ما إذا كانت تعرض وصفاً دقيقاً وكاملاً بشكل معقول للدراسة واستنتاجاتها. وأيضاً ينبغي أن يتأكد من عدم إدخال أية معلومات هنا لم يتم تضمينها في الأجزاء السابقة المناسبة. وثمة فكرة جيدة بأن يقوم زميل بقراءة جزء الاستنتاجات ليرى ما إذا كان الباحث يوصل القدر نفسه الذي نوى القيام به.

صفحات ملحقة Supplementary Pages

ثبت المراجع: يجب أن يتضمن ثبت المراجع جميع المصادر المذكورة في النص أو الهوامش. وتصر أغلب الجامعات على إدراج هذه فقط، غير أن بعضاً منها تطلب إدراج المراجع ذات الصلة من التي لم تذكر بشكل خاص. إن كتيب الأسلوب المختار مسبقا يعطي التفاصيل الكاملة حول طريقة إدراج المراجع. ومن المهم اتباع هذه القواعد بشكل صارم وتام. وفي الواقع، إنها استراتيجية جيدة أن نقوم بتعلمها قبل إجراء البحث من خلال أدبيات المقترح. وبإدراج كل مرجع بالشكل الصحيح حسبما تم مواجهته، يمكن للمرء أن يتجنب الوقت الإضافي لإيجاد المراجع مرة أخرى ووصفها بشكل كامل في ثبت المراجع. فمن المناسب إدراجها في بطاقات أو إدخالها في الحاسوب بحيث يمكن خزنها حسب الترتيب.

الملاحق: تحتوي الملاحق على المواد ذات الصلة التي لا تكون مهمة بما يكفي لتضمينها في صلب التقرير، لكنها قد تكون ذات قيمة لبعض القراء. وقد تتضمن مثل هذه المواد نسخاً

كاملة من الاختبارات المصممة محلياً أو الاستبيانات إضافة إلى التعليمات ومفاتيح وضع الدرجات لمثل هذه الأدوات وبيانات تحليل البنود للقياسات المستخدمة، والتعليمات الشفوية للأفراد، والجداول الطويلة أو ذات الأهمية الثانوية للدراسة.

السيرة: قد يطلب أحياناً من مؤلفي التقارير البحثية تضمين تقرير موجز حول تدريبهم، وخبرتهم، وعضويتهم المهنية، ومساهماتهم السابقة.

الخلاصة (abstract) (*): تطلب أغلب المؤسسات خلاصة وجيزة (abstract) منفصلة للأطروحة والتي ينبغي أن تتضمن صياغة محدودة بوضوح للمشكلة وأوصافا موجزة لطرق البحث، والنتائج الرئيسية والاستنتاجات / النتائج النهائية (Conclusions) والمضامين. ويجب أن تكون هذه الخلاصة محددة الطول (600 كلمة أو أقل). وترقم صفحات الخلاصة بشكل منفصل وتوضع في بداية أو نهاية الأطروحة.

مقال / مقالة المجلة The Journal Article

لدى إعداد مقال بحثي للنشر في مجلة، فإن الخطوة الأولى الجيدة هي النظر خلال ثبت مراجع المرء لتحديد أية مجلة قد نشرت اكبر مقدار من العمل في مجال اهتمامه. فالمعلومات المتعلقة بإجراء تقديم المخطوطات ستكون عادة موجودة داخل الغلاف الأمامي للمجلة. والعديد من المجلات تحدد أي كتيب خاص بالأسلوب سيستخدم (مثلاً، كتيب النشر لجمعية علم النفس الأمريكية، أو كتيب NEA للأسلوب). فإذا لم يتحدد أي كتيب، فإن الأسلوب المفضل، وطريقة كتابة المراجع، وما إلى ذلك يمكن تحديدها من دراسة المقالات المضمنة في العدد الأخير من المجلة.

يتبع المقال البحثي الخطة العامة ذاتها مثل الأطروحة، لكنه يجب أن يكون أقصر(**). فالرسالة أو الأطروحة تعمل على إثبات كفاءة الطالب وتتطلب إعدادا كاملا للبحث ذي العلاقة، ووصفاً كاملاً للإجراءات، وجدولة كاملة للنتائج (results)، واجتهاداً تأملياً. أما مقال

(*) تستخدم كلمتا (summary ،abstract) بمعنى "خلاصة" في اللغة العربية، بيد أن هناك فرقاً بينهما في البحث العلمي. فكلمة خلاصة (summary) تعني موجزاً (مطولاً)، يعرض فيه الباحث مختلف الأفكار والمفاهيم والطرق والاستنتاجات المهمة. أما الخلاصة الوجيزة (abstract) فإنها محدودة جداً وتعرض بإيجاز شديد ما تم في البحث حسبما تبين الفقرة أعلاه وقد تدعى خلاصة الخلاصة – (المراجع).

(**) ننوّه هنا إلى أن بعضهم يقترح أن لا تزيد الخلاصة الوجيزة (abstract) التي توضع في بداية المقالة، عن خمسة أو ستة جمل، ويقترح بعضهم أن لا تزيد عن 120 كلمة وآخرون يقترحون 180 كلمة على الأكثر. وتعرض المعلومات الآتية: 1) عبارة توضح الهدف من المشكلة (وتنطوي هذه على مشكلة البحث)، 2) وصف الأفراد المشاركين في الدراسة، 3) توضيح موجز لما قام به المشاركون أثناء الدراسة، 4) خلاصة بأهم الاستنتاجات والنتائج. ويستخدم بعض المؤلفين واحداً فقط من مصطلحي المناقشة (discussion) أو النتائج النهائية (conclusions)، ولم تجر العادة على استخدامهما معاً – ومع ذلك فإننا نجد أحياناً، بل كثيراً، من يتجاوزون هذه القواعد – (المراجع).

المجلة، من ناحية أخرى، فإنه يتطلب مجرد إيصال مساهمة المؤلف بالمعرفة. ولأجل الاقتصاد بسعة المجلة ووقت القراء، فإنه ينبغي أن يكون المقال مختصراً. فالجزء الخاص بالأدبيات يحتوي فقط على النتائج والأدلة التي توفر أساساً للمشكلة. وتقدم الصياغة العامة للمشكلة في فقرة واحدة، أو حتى أنها قد تحذف، وفي مثل هذه الحالة قد تبدأ المقالة بالفرضية. ويقدم جزء الإجراءات أيضاً بشكل مختصر جدا، رغم أن جميع المعلومات المطلوبة لإعادة / استنساخ التجربة ينبغي إدخالها إن أمكن ذلك. أما الجزء الخاص بالنتائج (results) فإنه ذو أهمية كبيرة لدى القارئ ومن ثم سيمثل جزءا من المقالة أكبر مما عليه الحال في الأطروحة. وينبغي فقط مناقشة أهم الاستنتاجات (findings) بشيء من التفصيل.

وينبغي أن ترفق رسالة التغطية / الغلاف المختصرة بالمخطوطة. ويبعث المحرر عادة للمؤلف ببطاقة بريدية يشكره على استلام المخطوطة ويقوم بتوزيع نسخ منها على أعضاء هيئة التحرير. ومن هذه النقطة، يمضي وقت كثير عادة قبل أن يعلم المؤلف ما إذا كانت المقالة قد قبلت (6 أسابيع هي مدة نموذجية). وبعد قبول المقالة فإن الأمر يستغرق عدة اشهر قبل نشرها.

وعندما ترفض مجلة المخطوطة فإنها ترسل إشعارا بالرفض مرفقاً بأسباب هذا الرفض. إن رفض مجلة واحدة لا يعني بالضرورة أن المقال غير جدير بالنشر. فثمة عدد من العوامل – مثل المنافسة على السعة، والتغيرات في سياسة التحرير، أو تحيز المراجعين – قد تؤثر على قرار النشر. فالمقال الذي رفضته مجلة واحدة قد ينقح ويقدم إلى مجلة أخرى. وتدور العديد من المقالات على عدة مجلات قبل أن تنشر. وعلى أية حال، فإنه ليس أخلاقياً تقديم المقال لأكثر من مجلة في آن واحد.

المؤتمر المهني The Professional Convention

يجد العديد من الباحثين بأن سماع أوراق بحثية، وهي تُقرأ في المؤتمرات المهنية هي طريقة جيدة لمواكبة تحديث الإطلاع في حقلهم. ويعود سبب ذلك إلى مرور فترة طويلة من الزمن بين إكمال مشروع البحث ونشره. وهذا التلكؤ الزمني طويل في الغالب بما جعل المجلات المهنية توصف أحياناً بكونها أرشيفية في طبيعتها.

وتعد الأوراق البحثية المعروضة في الاجتماعات بالطريقة نفسها كمقالات المجلات. وهي ليست بالضرورة تقارير بحث كامل، بل قد تكون تقارير تقدم في مشروعات جارية. وورقة البحث المقروءة أقل رسمية من مقال المجلة ويمكن تكييفها على نحو أدق لجمهورها. فبالوسع التوقع بأن الجمهور، على العموم، يألف تفاصيل البحث ذي الصلة وطرق القياس.

وتنظم عموماً كما يلي:

1- الصياغة المباشرة للفرضية،

2- وصف موجز للإجراءات،

3- النتائج والاستنتاجات والمضامين.

أن الوقت المسموح لقراءة ورقة بحثيه هو وقت مختصر في العادة، ويقل عن 30 دقيقة. وعليه ينبغي أن تركز الورقة البحثية على أهم جوانب الدراسة. والقاعدة العملية المناسبة هي ½ 2 – 3 دقيقة لكل صفحة من النسخة المطبوعة بفراغ مضاعف.

فإذا كانت الأشكال والجداول ستساعد في التقديم فيجب أن تتوفر نسخ منها للجمهور. ويوزع بعض المتحدثين نسخاً من نص ورقتهم البحثية على الجمهور. وعلى أية حال، بما أن الجمهور يقرأ الورقة البحثية بصمت بأسرع مما يستطيع أن يقراه المؤلف بصوت عال فقد تكون هناك ثغرة في الانتباه. ومن المفضل أن تكون هناك عدة نسخ من الوصف الكامل للدراسة متيسرة لكل من يطلبها بعد تقديم الورقة.

قائمة تدقيق/ مراجعة لتقييم التقارير البحثية

CHECKLIST FOR EVALUATION RESEARCH REPORTS

لابد أن تكون قائمة التدقيق/ المراجعة مفيدة في تقييم تقارير بحوثك وبحوث الآخرين. وهي تجمع معا العديد من الموضوعات المقدمة في هذا الكتاب.

العنوان	هل يوضح العنوان المجتمع الإحصائي المعني والمتغيرات الرئيسية؟ هل تم تجنب المصطلحات الغامضة، والملتبسة، والمشحونة بالعواطف؟
صياغة المشكلة	هل تم تحديد المتغيرات المعنية؟ هل تم تطوير الأساس النظري للدراسة؟ هل تم صياغة المشكلة بوضوح؟
الأدبيات ذات الصلة	هل الأدبيات ذات الصلة وكافية؟ هل الارتباطات بين الدراسة الحالية والبحوث السابقة والنظرية واضحة؟
الفرضيات	هل الفرضيات واضحة؟ هل تنبع الفرضيات بشكل منطقي عبارة المشكلة؟
الأفراد	هل المجتمع الإحصائي المعنى محدد؟ هل طريقة انتخاب العينة واضحة؟ هل تسمح المعاينة بالتعميم على المجتمع الإحصائي المعني؟
الإجراءات	هل الإجراءات موصوفة بشكل مناسب يكفي بالسماح لإعادة/ استنساخ الدراسة؟ هل تتضمن تعريفات إجرائية للمتغيرات المستقلة؟ هل توفر الإجراءات ضبطا كافيا للصدق الداخلي؟ هل توفر الإجراءات ضبطا كافيا للصدق الخارجي؟

الأدوات	هل الأدوات موصوفة بشكل كاف؟ هل المعلومات حول صدق وثبات الأدوات متوفرة؟ هل الأدوات، تعريفات إجرائية مناسبة للمتغيرات التابعة؟
تحليل البيانات	هل الإحصاء الوصفي المستخدم مناسب لتلخيص البيانات؟ هل الإحصاء الاستدلالي المستخدم مناسب لاختبار الفرضيات؟ هل فسرت الإحصائيات بشكل مناسب؟ هل الإحصائيات مناسبة لمستوى قياس البيانات؟
النتائج Results	هل عرضت نتائج جميع اختبارات الفرضيات؟ هل تم تفسير الإحصاء بشكل صحيح؟ هل النتائج معروضة بشكل صحيح؟ هل الصلات بين النتائج والنظرية والأدبيات ذات العلاقة واضحة؟
النتائج النهائية Conclusions	هل الاستنتاجات معروضة بشكل واضح؟ هل تنبع الاستنتاجات، بشكل منطقي، نتائج الدراسة؟ هل تجنب المؤلف التوصل إلى استنتاجات لا تدعمها بشكل مباشر نتائج الدراسة؟
المضامين والتطبيقات	هل يعرض المؤلف مضامين نتائج الدراسة؟ هل المضامين المعروضة مستندة إلى نتائج الدراسة، وليس إلى ما يأمله أو يتوقعه المؤلف صحيحا؟ هل نوقشت التطبيقات المناسبة؟ هل تنبع التطبيقات بشكل منطقي نتائج الدراسة؟
الخلاصة Summary	هل الخلاصة واضحة ودقيقة وكاملة بشكل كاف؟

كتيبات الأسلوب STYLE MANUALS

فيما يلي كتيبات / كراسات واسعة الاستخدام تفصل الشكل العام والأسلوب للرسائل والأطروحات:

American Psychological Association. (1994). *Publication manual* (4th ed.). Washington, DC: American Psychological Association.

Campbell, W.G., and Ballou, s.v. (1990). *Form and style: Theses, reports, term papers* (8thed.). Boston: Houghton Mifflin.

University of Chicago Press. (1993). *The Chicago manual of style* (14th ed.). Chicago: University of Chicago Press.

Turabian, K. (1987). *A manual for writers of term papers, theses, and dissertations* (5th ed.). Chicago: University of Chicago Press.

الخلاصة SUMMARY

إن استخدام الحاسوب يوفر الوقت والطاقة ويسهل درجة عالية من الدقة في حسابات معقدة جداً. ويحتاج الباحث لدى استخدام الحاسوب لمعالجة البيانات إلى تحويل البيانات إلى شكل يمكن قراءته في الحاسوب وتزويد الحاسوب بالتعليمات للعمل على البيانات عن طريق برنامج مناسب. إن كتابة برامج الحاسوب تتطلب التدريب، إلا أنه بوسع الباحث عادة أن يجد برامج متوفرة معدة للاستخدام في تحليل البيانات.

إن تفسير نتائج الدراسة عمل مباشر إذا وضع الباحث ، في مقترح البحث، أساساً مناسباً للدراسة البحثية. وينبغي أن لا تغيب عنا التحذيرات التالية: 1) ينبغي أن يستند التفسير إلى البيانات المستمدة من الدراسة، 2) ينبغي النظر في مشكلات الصدق الداخلي والخارجي والحدود الأخرى للدراسة، 3) يجب عرض الاستنتاجات كعبارات احتمالية وليس كحقائق.

تستحق النتائج السالبة التقدير ذاته والتفسير كما هو الحال مع النتائج الموجبة. وتفسر الفرضية الصفرية المحتفظ بها كمؤشر لدليل غير كاف ولا أكثر. وتستحق النتائج غير المفترضة الاهتمام كمصادر لفرضيات مستقبلية.

لقد تم تطوير إجراءات شكلية(*) لإعداد أطروحات، ومقالات، مجلات، وأوراق بحث. إن إتقان هذه الإجراءات واستخدامها يساعد الباحثين على الاتصال مع بعضهم ومع التربويين الممارسين. وثمة العديد من كتيبات/ كراسات الأسلوب توفر تفصيلات حول هذه الإجراءات

مفاهيم أساسية Key Concepts

interpreting expected results	تفسير النتائج المتوقعة
interpreting negative results	تفسير النتائج السالبة
interpreting retained null hypotheses	تفسير الفرضيات الصفرية المحتفظ بها/ المستبقاة
interpreting unexpected results	تفسير النتائج غير المتوقعة
practical versus statistical significance	الأهمية العملية مقابل الدلالة الإحصائية
preparing conference papers	إعداد أوراق بحث للمؤتمرات
preparing journal articles	إعداد مقالات المجلات
preparing theses or dissertations	إعداد الرسائل أو الأطروحات
style manuals	كتيبات الأسلوب

(*) شكلية / صورية / رسمية (formal): تشير إلى مجموعة من القواعد التي تضبط موقفا ما بطريقة نظامية ومنطقية ومخططة ومنهجية – (المراجع).

تمارين EXERCIES

1- ما هي التغيرات التي جاء بها الحاسوب في ممارسة البحث التربوي؟

2- حدد ثلاثة برامج حاسوبية متوفرة تجاريا للتحليل الإحصائي؟

3- ما هو الفرق بين النتائج (results) والاستنتاجات (conclusions)؟

4- وضح الفرق بين الدلالة (significance) الإحصائية للنتائج وأهمية significance الدراسة.

5- ما هي الحالات التي يمكن أن تؤدي إلى الاحتفاظ بالفرضية الصفرية؟

6- قرر ما إذا كان كل من العناوين التالية مقبولا أو غير مقبول وأعط أسبابا لاختيارك:

أ. معدل نقاط (GPA) التخرج وتعليم السياقة

ب. آثار التدريس الفردي الذي يقوم به طلبة الصف السادس، بمعدل ثلاث مرات في الأسبوع على أداء القراءة لتلاميذ الصف الثاني الذين مستواهم القرائي دون المعدل.

ج. ينبغي تدريس الأطفال الرياضيات الحديثة.

د. العلاقة بين سمات الشخصية والاتجاهات نحو تحصيل الطلبة الجيدين والضعفاء في القراءة.

7- ما هي الفروق في الشكل العام للبحث المكتوب في صيغة أطروحة، أو مجلة أو ورقة بحث تقرأ في المؤتمر؟

8- هل ينبغي على المرء أن يناقش نتائج البحث التي لا تتفق مع فرضياته؟

الأجوبة ANSWERS

1- لقد جعل الحاسوب بالإمكان تصميم وإجراء الدراسات البحثية دون اعتبار لعدد المتغيرات أو لتعقيد التحليل. والبيانات من الدراسات التي فيها أعداد كبيرة من الأفراد، أو استخدام الأساليب متعددة المتغيرات، أو التي تتطلب الإحصائيات المعقدة يمكن الآن تحليلها بسرعة ودقة.

2- أي ثلاثة من بين برامج MYSTAT، SPSS، student SPSS، SPSS–Pc، SYSTAT، BMDP، SAS، STATGRAPH.

3- النتيجة (result) هي ملاحظة مباشرة. والاستنتاج (conclusion) هو استدلال يستند إلى النتائج (results).

4- الدلالة significance الإحصائية تعني فقط ان النتائج لا يحتمل أن تكون بفعل صدفة، أما أهمية (significance) الدراسة فتتحدد بأهمية النتائج فيما يخص اختبار نظرية أو مضامين عملية.

5- يمكن أن تأتي الفرضية الصفرية المستبقاة من كون الفرضية الصفرية صحيحة في طبيعتها، أو قد تنتج عن التشويش / الإفساد الذي تخلفه مشكلات الصدق الداخلي بما يحجب آثار المعالجة، ومن الافتقار للقوة الإحصائية للتصميم المستخدم في الدراسة، أو من عجز الأدوات في القياس الدقيق لآثار المعالجة على المتغير التابع.

6- أ. غير مقبول: لا تتوفر عبارة للعلاقة.
ب. غير مقبول: كلام كثير.
ج. غير مقبول: عناوين مشحونة بالعواطف فهي غير مناسبة للمقالات البحثية.
د. مقبول: يلبي معايير العنوان.

7- شكل الأطروحة هو العرض الأكثر تفصيلا والأكثر شكلية / رسمية. وهو يتبع تفاصيل كتيب الأسلوب الخاص. أما مقال المجلة فهو عرض اكثر إيجازاً، مع بيان موجز للمشكلة وللأدبيات ذات الصلة والمنهجية، ونسبة أكبر من المقالة مكرسة للنتائج (results) الرئيسية ومناقشة أهميتها. وورقة البحث تعد لتقرأ في مؤتمر هو الأكثر ابتعاداً عن الرسمية ومتكيّفيه مع جمهورها، وهي تعرض الفرضية، وتصف الإجراء بإيجاز وتؤكد على أهم الاستنتاجات (findings).

8- نعم أن النتائج (results) المخالفة لتوقعات الفرد هي نتائج مشروعة مثل أية نتائج أخرى ويجب تفسيرها في حد ذاتها.

المصادر REFERENCES

Bowker. R. R. (1993). The software encyclopedia. New York: R.R. Bowker.

Appendix الملحق

Table A.1 Areas of the Normal Curve

(1) z	(2) Area between the Mean and z	(3) Area beyond z	(1) z	(2) Area between the Mean and z	(3) Area beyond z
0.00	.0000	.5000	0.35	.1368	.3632
0.01	.0040	.4960	0.36	.1406	.3594
0.02	.0080	.4920	0.37	.1443	.3557
0.03	.0120	.4880	0.38	.1480	.3520
0.04	.0160	.4840	0.39	.1517	.3483
0.05	.0199	.4801	0.40	.1554	.3446
0.06	.0239	.4761	0.41	.1591	.3409
0.07	.0279	.4721	0.42	.1628	.3372
0.08	.0319	.4681	0.43	.1664	.3336
0.09	.0359	.4641	0.44	.1700	.3300
0.10	.0398	.4602	0.45	.1736	.3264
0.11	.0438	.4562	0.46	.1772	.3228
0.12	.0478	.4522	0.47	.1808	.3192
0.13	.0517	.4483	0.48	.1844	.3156
0.14	.0557	.4443	0.49	.1879	.3121
0.15	.0596	.4404	0.50	.1915	.3085
0.16	.0636	.4364	0.51	.1950	.3050
0.17	.0675	.4325	0.52	.1985	.3015
0.18	.0714	.4286	0.53	.2019	.2981
0.19	.0753	.4247	0.54	.2054	.2946
0.20	.0793	.4207	0.55	.2088	.2912
0.21	.0832	.4168	0.56	.2123	.2877
0.22	.0871	.4129	0.57	.2157	.2843
0.23	.0910	.4090	0.58	.2190	.2810
0.24	.0948	.4052	0.59	.2224	.2776
0.25	.0987	.4013	0.60	.2257	.2743
0.26	.1026	.3974	0.61	.2291	.2709
0.27	.1064	.3936	0.62	.2324	.2676
0.28	.1103	.3897	0.63	.2357	.2643
0.29	.1141	.3859	0.64	.2389	.2611
0.30	.1179	.3821	0.65	.2422	.2578
0.31	.1217	.3783	0.66	.2454	.2546
0.32	.1255	.3745	0.67	.2486	.2514
0.33	.1293	.3707	0.68	.2517	.2483
0.34	.1331	.3669	0.69	.2549	.2451

Table A.1 Areas of the Normal Curve (*cont.*)

(1) z	(2) Area between the Mean and z	(3) Area beyond z	(1) z	(2) Area between the Mean and z	(3) Area beyond z
0.70	.2580	.2420	1.05	.3531	.1469
0.71	.2611	.2389	1.06	.3554	.1446
0.72	.2642	.2358	1.07	.3577	.1423
0.73	.2673	.2327	1.08	.3599	.1401
0.74	.2704	.2296	1.09	.3621	.1379
0.75	.2734	.2266	1.10	.3643	.1357
0.76	.2764	.2236	1.11	.3665	.1335
0.77	.2794	.2206	1.12	.3686	.1314
0.78	.2823	.2177	1.13	.3708	.1292
0.79	.2852	.2148	1.14	.3729	.1271
0.80	.2881	.2119	1.15	.3749	.1251
0.81	.2910	.2090	1.16	.3770	.1230
0.82	.2939	.2061	1.17	.3790	.1210
0.83	.2967	.2033	1.18	.3810	.1190
0.84	.2995	.2005	1.19	.3830	.1170
0.85	.3023	.1977	1.20	.3849	.1151
0.86	.3051	.1949	1.21	.3869	.1131
0.87	.3078	.1922	1.22	.3888	.1112
0.88	.3106	.1894	1.23	.3907	.1093
0.89	.3133	.1867	1.24	.3925	.1075
0.90	.3159	.1841	1.25	.3944	.1056
0.91	.3186	.1814	1.26	.3962	.1038
0.92	.3212	.1788	1.27	.3980	.1020
0.93	.3238	.1762	1.28	.3997	.1003
0.94	.3264	.1736	1.29	.4015	.0985
0.95	.3289	.1711	1.30	.4032	.0968
0.96	.3315	.1685	1.31	.4049	.0951
0.97	.3340	.1660	1.32	.4066	.0934
0.98	.3365	.1635	1.33	.4082	.0918
0.99	.3389	.1611	1.34	.4099	.0901
1.00	.3413	.1587	1.35	.4115	.0885
1.01	.3438	.1562	1.36	.4131	.0869
1.02	.3461	.1539	1.37	.4147	.0853
1.03	.3485	.1515	1.38	.4162	.0838
1.04	.3508	.1492	1.39	.4177	.0823

Table A.1 Areas of the Normal Curve (*cont.*)

(1) z	(2) Area between the Mean and z	(3) Area beyond z	(1) z	(2) Area between the Mean and z	(3) Area beyond z
1.40	.4192	.0808	1.75	.4599	.0401
1.41	.4207	.0793	1.76	.4608	.0392
1.42	.4222	.0778	1.77	.4616	.0384
1.43	.4236	.0764	1.78	.4625	.0375
1.44	.4251	.0749	1.79	.4633	.0367
1.45	.4265	.0735	1.80	.4641	.0359
1.46	.4279	.0721	1.81	.4649	.0351
1.47	.4292	.0708	1.82	.4656	.0344
1.48	.4306	.0694	1.83	.4664	.0336
1.49	.4319	.0681	1.84	.4671	.0329
1.50	.4332	.0668	1.85	.4678	.0322
1.51	.4345	.0655	1.86	.4686	.0314
1.52	.4357	.0643	1.87	.4693	.0307
1.53	.4370	.0630	1.88	.4699	.0301
1.54	.4382	.0618	1.89	.4706	.0294
1.55	.4394	.0606	1.90	.4713	.0287
1.56	.4406	.0594	1.91	.4719	.0281
1.57	.4418	.0582	1.92	.4726	.0274
1.58	.4429	.0571	1.93	.4732	.0268
1.59	.4441	.0559	1.94	.4738	.0262
1.60	.4452	.0548	1.95	.4744	.0256
1.61	.4463	.0537	1.96	.4750	.0250
1.62	.4474	.0526	1.97	.4756	.0244
1.63	.4484	.0516	1.98	.4761	.0239
1.64	.4495	.0505	1.99	.4767	.0233
1.65	.4505	.0495	2.00	.4772	.0228
1.66	.4515	.0485	2.01	.4778	.0222
1.67	.4525	.0475	2.02	.4783	.0217
1.68	.4535	.0465	2.03	.4788	.0212
1.69	.4545	.0455	2.04	.4793	.0207
1.70	.4554	.0446	2.05	.4798	.0202
1.71	.4564	.0436	2.06	.4803	.0197
1.72	.4573	.0427	2.07	.4808	.0192
1.73	.4582	.0418	2.08	.4812	.0188
1.74	.4591	.0409	2.09	.4817	.0183

Table A.1 Areas of the Normal Curve (*cont.*)

(1) z	(2) Area between the Mean and z	(3) Area beyond z	(1) z	(2) Area between the Mean and z	(3) Area beyond z
2.10	.4821	.0179	2.45	.4929	.0071
2.11	.4826	.0174	2.46	.4931	.0069
2.12	.4830	.0170	2.47	.4932	.0068
2.13	.4834	.0166	2.48	.4934	.0066
2.14	.4838	.0162	2.49	.4936	.0064
2.15	.4842	.0158	2.50	.4938	.0062
2.16	.4846	.0154	2.51	.4940	.0060
2.17	.4850	.0150	2.52	.4941	.0059
2.18	.4854	.0146	2.53	.4943	.0057
2.19	.4857	.0143	2.54	.4945	.0055
2.20	.4861	.0139	2.55	.4946	.0054
2.21	.4864	.0136	2.56	.4948	.0052
2.22	.4868	.0132	2.57	.4949	.0051
2.23	.4871	.0129	2.58	.4951	.0049
2.24	.4875	.0125	2.59	.4952	.0048
2.25	.4878	.0122	2.60	.4953	.0047
2.26	.4881	.0119	2.61	.4955	.0045
2.27	.4884	.0116	2.62	.4956	.0044
2.28	.4887	.0113	2.63	.4957	.0043
2.29	.4890	.0110	2.64	.4959	.0041
2.30	.4893	.0107	2.65	.4960	.0040
2.31	.4896	.0104	2.66	.4961	.0039
2.32	.4898	.0102	2.67	.4962	.0038
2.33	.4901	.0099	2.68	.4963	.0037
2.34	.4904	.0096	2.69	.4964	.0036
2.35	.4906	.0094	2.70	.4965	.0035
2.36	.4909	.0091	2.71	.4966	.0034
2.37	.4911	.0089	2.72	.4967	.0033
2.38	.4913	.0087	2.73	.4968	.0032
2.39	.4916	.0084	2.74	.4969	.0031
2.40	.4918	.0082	2.75	.4970	.0030
2.41	.4920	.0080	2.76	.4971	.0029
2.42	.4922	.0078	2.77	.4972	.0028
2.43	.4925	.0075	2.78	.4973	.0027
2.44	.4927	.0073	2.79	.4974	.0026

Table A.1 Areas of the Normal Curve (*cont.*)

(1) z	(2) Area between the Mean and z	(3) Area beyond z	(1) z	(2) Area between the Mean and z	(3) Area beyond z
2.80	.4974	.0026	3.10	.4990	.0010
2.81	.4975	.0025	3.11	.4991	.0009
2.82	.4976	.0023	3.12	.4991	.0009
2.83	.4977	.0024	3.13	.4991	.0009
2.84	.4977	.0023	3.14	.4992	.0008
2.85	.4978	.0022	3.15	.4992	.0008
2.86	.4979	.0021	3.16	.4992	.0008
2.87	.4979	.0021	3.17	.4992	.0008
2.88	.4980	.0020	3.18	.4993	.0007
2.89	.4981	.0019	3.19	.4993	.0007
2.90	.4981	.0019	3.20	.4993	.0007
2.91	.4982	.0018	3.21	.4993	.0007
2.92	.4982	.0018	3.22	.4994	.0006
2.93	.4983	.0017	3.23	.4994	.0006
2.94	.4984	.0016	3.24	.4994	.0006
2.95	.4984	.0016	3.30	.4995	.0005
2.96	.4985	.0015	3.40	.4997	.0003
2.97	.4985	.0015	3.50	.4998	.0002
2.98	.4986	.0014	3.60	.4998	.0002
2.99	.4986	.0014	3.70	.4999	.0001
3.00	.4987	.0013	3.90	.49995	.00005
3.01	.4987	.0013	4.00	.49997	.00003
3.02	.4987	.0013	4.50	.4999966	.0000034
3.03	.4988	.0012	5.00	.4999997	.0000003
3.04	.4988	.0012	5.50	.499999981	.000000019
3.05	.4989	.0011	6.00	.499999999	.000000001
3.06	.4989	.0011			
3.07	.4989	.0011			
3.08	.4990	.0010			
3.09	.4990	.0010			

Table A.2 Table of *t*-Values

	Level of significance for a directional (one-tailed) test					
	.10	.05	.025	.01	.005	.0005
	Level of significance for a nondirectional (two-tailed) test					
df	.20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

Source: Table A.2 is taken from Table III of Fisher and Yates, *Statistical Tables for Biological, Agricultural, and Medical Research*, published by Longman Group UK Ltd., 1974.

Table A.3 The 5 (Roman Type) and 1 (Boldface Type) Percent Points for the F Distribution

denominator $df = n_2$	n_1 degrees of freedom for numerator																							
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	253	254	254	254
	4,052	**4,999**	**5,403**	**5,625**	**5,764**	**5,859**	**5,928**	**5,981**	**6,022**	**6,056**	**6,082**	**6,106**	**6,142**	**6,169**	**6,208**	**6,234**	**6,258**	**6,286**	**6,302**	**6,323**	**6,334**	**6,352**	**6,361**	**6,366**
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41	19.42	19.43	19.44	19.45	19.46	19.47	19.47	19.48	19.49	19.49	19.50	19.50
	98.49	**99.00**	**99.17**	**99.25**	**99.30**	**99.33**	**99.34**	**99.36**	**99.38**	**99.40**	**99.41**	**99.42**	**99.43**	**99.44**	**99.45**	**99.46**	**99.47**	**99.48**	**99.48**	**99.49**	**99.49**	**99.49**	**99.50**	**99.50**
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74	8.71	8.69	8.66	8.64	8.62	8.60	8.58	8.57	8.56	8.54	8.54	8.53
	34.12	**30.82**	**29.46**	**28.71**	**28.24**	**27.91**	**27.67**	**27.49**	**27.34**	**27.23**	**27.13**	**27.05**	**26.92**	**26.83**	**26.69**	**26.60**	**26.50**	**26.41**	**26.35**	**26.27**	**26.23**	**26.18**	**26.14**	**26.12**
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91	5.87	5.84	5.80	5.77	5.74	5.71	5.70	5.68	5.66	5.65	5.64	5.63
	21.20	**18.00**	**16.69**	**15.98**	**15.52**	**15.21**	**14.98**	**14.80**	**14.66**	**14.54**	**14.45**	**14.37**	**14.24**	**14.15**	**14.02**	**13.93**	**13.83**	**13.74**	**13.69**	**13.61**	**13.57**	**13.52**	**13.48**	**13.46**
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.40	4.38	4.37	4.36
	16.26	**13.27**	**12.06**	**11.39**	**10.97**	**10.67**	**10.45**	**10.27**	**10.15**	**10.05**	**9.96**	**9.89**	**9.77**	**9.68**	**9.55**	**9.47**	**9.38**	**9.29**	**9.24**	**9.17**	**9.13**	**9.07**	**9.04**	**9.02**
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68	3.67
	13.74	**10.92**	**9.78**	**9.15**	**8.75**	**8.47**	**8.26**	**8.10**	**7.98**	**7.87**	**7.79**	**7.72**	**7.60**	**7.52**	**7.39**	**7.31**	**7.23**	**7.14**	**7.09**	**7.02**	**6.99**	**6.94**	**6.90**	**6.88**
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24	3.23
	12.25	**9.55**	**8.45**	**7.85**	**7.46**	**7.19**	**7.00**	**6.84**	**6.71**	**6.62**	**6.54**	**6.47**	**6.35**	**6.27**	**6.15**	**6.07**	**5.98**	**5.90**	**5.85**	**5.78**	**5.75**	**5.70**	**5.67**	**5.65**
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.94	2.93
	11.26	**8.65**	**7.59**	**7.01**	**6.63**	**6.37**	**6.19**	**6.03**	**5.91**	**5.82**	**5.74**	**5.67**	**5.56**	**5.48**	**5.36**	**5.28**	**5.20**	**5.11**	**5.06**	**5.00**	**4.96**	**4.91**	**4.88**	**4.86**
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.76	2.73	2.72	2.71
	10.56	**8.02**	**6.99**	**6.42**	**6.06**	**5.80**	**5.62**	**5.47**	**5.35**	**5.26**	**5.18**	**5.11**	**5.00**	**4.92**	**4.80**	**4.73**	**4.64**	**4.56**	**4.51**	**4.45**	**4.41**	**4.36**	**4.33**	**4.31**
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.55	2.54
	10.04	**7.56**	**6.55**	**5.99**	**5.64**	**5.39**	**5.21**	**5.06**	**4.95**	**4.85**	**4.78**	**4.71**	**4.60**	**4.52**	**4.41**	**4.33**	**4.25**	**4.17**	**4.12**	**4.05**	**4.01**	**3.96**	**3.93**	**3.91**
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.41	2.40
	9.65	**7.20**	**6.22**	**5.67**	**5.32**	**5.07**	**4.88**	**4.74**	**4.63**	**4.54**	**4.46**	**4.40**	**4.29**	**4.21**	**4.10**	**4.02**	**3.94**	**3.86**	**3.80**	**3.74**	**3.70**	**3.66**	**3.62**	**3.60**
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.31	2.30
	9.33	**6.93**	**5.95**	**5.41**	**5.06**	**4.82**	**4.65**	**4.50**	**4.39**	**4.30**	**4.22**	**4.16**	**4.05**	**3.98**	**3.86**	**3.78**	**3.70**	**3.61**	**3.56**	**3.49**	**3.46**	**3.41**	**3.38**	**3.36**
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60	2.55	2.51	2.46	2.42	2.38	2.34	2.32	2.28	2.26	2.24	2.22	2.21
	9.07	**6.70**	**5.74**	**5.20**	**4.86**	**4.62**	**4.44**	**4.30**	**4.19**	**4.10**	**4.02**	**3.96**	**3.85**	**3.78**	**3.67**	**3.59**	**3.51**	**3.42**	**3.37**	**3.30**	**3.27**	**3.21**	**3.18**	**3.16**

Table A.3 The 5 (Roman Type) and 1 (Boldface Type) Percent Points for the *F* Distribution (*conf.*)

denominator	n_1 degrees of freedom for numerator																							
$df = n_1$	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.16	2.14	2.13
	8.86	**6.51**	**5.56**	**5.03**	**4.69**	**4.46**	**4.28**	**4.14**	**4.03**	**3.94**	**3.86**	**3.80**	**3.70**	**3.62**	**3.51**	**3.43**	**3.34**	**3.26**	**3.21**	**3.14**	**3.11**	**3.06**	**3.02**	**3.00**
15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48	2.43	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.12	2.10	2.08	2.07
	8.68	**6.36**	**5.42**	**4.89**	**4.56**	**4.32**	**4.14**	**4.00**	**3.89**	**3.80**	**3.73**	**3.67**	**3.56**	**3.48**	**3.36**	**3.29**	**3.20**	**3.12**	**3.07**	**3.00**	**2.97**	**2.92**	**2.89**	**2.87**
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42	2.37	2.33	2.28	2.24	2.20	2.16	2.13	2.09	2.07	2.04	2.02	2.01
	8.53	**6.23**	**5.29**	**4.77**	**4.44**	**4.20**	**4.03**	**3.89**	**3.78**	**3.69**	**3.61**	**3.55**	**3.45**	**3.37**	**3.25**	**3.18**	**3.10**	**3.01**	**2.96**	**2.89**	**2.86**	**2.80**	**2.77**	**2.75**
17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.38	2.33	2.29	2.23	2.19	2.15	2.11	2.08	2.04	2.02	1.99	1.97	1.96
	8.40	**6.11**	**5.18**	**4.67**	**4.34**	**4.10**	**3.93**	**3.79**	**3.68**	**3.59**	**3.52**	**3.45**	**3.35**	**3.27**	**3.16**	**3.08**	**3.00**	**2.92**	**2.86**	**2.79**	**2.76**	**2.70**	**2.67**	**2.65**
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.29	2.25	2.19	2.15	2.11	2.07	2.04	2.00	1.98	1.95	1.93	1.92
	8.28	**6.01**	**5.09**	**4.58**	**4.25**	**4.01**	**3.85**	**3.71**	**3.60**	**3.51**	**3.44**	**3.37**	**3.27**	**3.19**	**3.07**	**3.00**	**2.91**	**2.83**	**2.78**	**2.71**	**2.68**	**2.62**	**2.59**	**2.57**
19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.31	2.26	2.21	2.15	2.11	2.07	2.02	2.00	1.96	1.94	1.91	1.90	1.88
	8.18	**5.93**	**5.01**	**4.50**	**4.17**	**3.94**	**3.77**	**3.63**	**3.52**	**3.43**	**3.36**	**3.30**	**3.19**	**3.12**	**3.00**	**2.92**	**2.84**	**2.76**	**2.70**	**2.63**	**2.60**	**2.54**	**2.51**	**2.49**
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.28	2.23	2.18	2.12	2.08	2.04	1.99	1.96	1.92	1.90	1.87	1.85	1.84
	8.10	**5.85**	**4.94**	**4.43**	**4.10**	**3.87**	**3.71**	**3.56**	**3.45**	**3.37**	**3.30**	**3.23**	**3.13**	**3.05**	**2.94**	**2.86**	**2.77**	**2.69**	**2.63**	**2.56**	**2.53**	**2.47**	**2.44**	**2.42**
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.20	2.15	2.09	2.05	2.00	1.96	1.93	1.89	1.87	1.84	1.82	1.81
	8.02	**5.78**	**4.87**	**4.37**	**4.04**	**3.81**	**3.65**	**3.51**	**3.40**	**3.31**	**3.24**	**3.17**	**3.07**	**2.99**	**2.88**	**2.80**	**2.72**	**2.63**	**2.58**	**2.51**	**2.47**	**2.42**	**2.38**	**2.36**
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.26	2.23	2.18	2.13	2.07	2.03	1.98	1.93	1.91	1.87	1.84	1.81	1.80	1.78
	7.94	**5.72**	**4.82**	**4.31**	**3.99**	**3.76**	**3.59**	**3.45**	**3.35**	**3.26**	**3.18**	**3.12**	**3.02**	**2.94**	**2.83**	**2.75**	**2.67**	**2.58**	**2.53**	**2.46**	**2.42**	**2.37**	**2.33**	**2.31**
23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32	2.28	2.24	2.20	2.14	2.10	2.04	2.00	1.96	1.91	1.88	1.84	1.82	1.79	1.77	1.76
	7.88	**5.66**	**4.76**	**4.26**	**3.94**	**3.71**	**3.54**	**3.41**	**3.30**	**3.21**	**3.14**	**3.07**	**2.97**	**2.89**	**2.78**	**2.70**	**2.62**	**2.53**	**2.48**	**2.41**	**2.37**	**2.32**	**2.28**	**2.26**
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.22	2.18	2.13	2.09	2.02	1.98	1.94	1.89	1.86	1.82	1.80	1.76	1.74	1.73
	7.82	**5.61**	**4.72**	**4.22**	**3.90**	**3.67**	**3.50**	**3.36**	**3.25**	**3.17**	**3.09**	**3.03**	**2.93**	**2.85**	**2.74**	**2.66**	**2.58**	**2.49**	**2.44**	**2.36**	**2.33**	**2.27**	**2.23**	**2.21**
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24	2.20	2.16	2.11	2.06	2.00	1.96	1.92	1.87	1.84	1.80	1.77	1.74	1.72	1.71
	7.77	**5.57**	**4.68**	**4.18**	**3.86**	**3.63**	**3.46**	**3.32**	**3.21**	**3.13**	**3.05**	**2.99**	**2.89**	**2.81**	**2.70**	**2.62**	**2.54**	**2.45**	**2.40**	**2.32**	**2.29**	**2.23**	**2.19**	**2.17**
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.10	2.05	1.99	1.95	1.90	1.85	1.82	1.78	1.76	1.72	1.70	1.69
	7.72	**5.53**	**4.64**	**4.14**	**3.82**	**3.59**	**3.42**	**3.29**	**3.17**	**3.09**	**3.02**	**2.96**	**2.86**	**2.77**	**2.66**	**2.58**	**2.50**	**2.41**	**2.36**	**2.28**	**2.25**	**2.19**	**2.15**	**2.13**

Table A.3 The 5 (Roman Type) and 1 (Boldface Type) Percent Points for the F Distribution (*cont.*)

denominator $df = n_1$	n_1 degrees of freedom for numerator: 1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20	2.16	2.13	2.08	2.03	1.97	1.93	1.88	1.84	1.80	1.76	1.74	1.71	1.68	1.67
	7.68	**5.49**	**4.60**	**4.11**	**3.79**	**3.56**	**3.39**	**3.26**	**3.14**	**3.06**	**2.98**	**2.93**	**2.83**	**2.74**	**2.63**	**2.55**	**2.47**	**2.38**	**2.33**	**2.25**	**2.21**	**2.16**	**2.12**	**2.10**
28	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24	2.19	2.15	2.12	2.06	2.02	1.96	1.91	1.87	1.81	1.78	1.75	1.72	1.69	1.67	1.65
	7.64	**5.45**	**4.57**	**4.07**	**3.76**	**3.53**	**3.36**	**3.23**	**3.11**	**3.03**	**2.95**	**2.90**	**2.80**	**2.71**	**2.60**	**2.52**	**2.44**	**2.35**	**2.30**	**2.22**	**2.18**	**2.13**	**2.09**	**2.06**
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.05	2.00	1.94	1.90	1.85	1.80	1.77	1.73	1.71	1.68	1.65	1.64
	7.60	**5.42**	**4.54**	**4.04**	**3.73**	**3.50**	**3.33**	**3.20**	**3.08**	**3.00**	**2.92**	**2.87**	**2.77**	**2.68**	**2.57**	**2.49**	**2.41**	**2.32**	**2.27**	**2.19**	**2.15**	**2.10**	**2.06**	**2.03**
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16	2.12	2.09	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.69	1.66	1.64	1.62
	7.56	**5.39**	**4.51**	**4.02**	**3.70**	**3.47**	**3.30**	**3.17**	**3.06**	**2.98**	**2.90**	**2.84**	**2.74**	**2.66**	**2.55**	**2.47**	**2.38**	**2.29**	**2.24**	**2.16**	**2.13**	**2.07**	**2.03**	**2.01**
32	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14	2.10	2.07	2.02	1.97	1.91	1.86	1.82	1.76	1.74	1.69	1.67	1.64	1.61	1.59
	7.50	**5.34**	**4.46**	**3.97**	**3.66**	**3.42**	**3.25**	**3.12**	**3.01**	**2.94**	**2.86**	**2.80**	**2.70**	**2.62**	**2.51**	**2.42**	**2.34**	**2.25**	**2.20**	**2.12**	**2.08**	**2.02**	**1.98**	**1.96**
34	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12	2.08	2.05	2.00	1.95	1.89	1.84	1.80	1.74	1.71	1.67	1.64	1.61	1.59	1.57
	7.44	**5.29**	**4.42**	**3.93**	**3.61**	**3.38**	**3.21**	**3.08**	**2.97**	**2.89**	**2.82**	**2.76**	**2.66**	**2.58**	**2.47**	**2.38**	**2.30**	**2.21**	**2.15**	**2.08**	**2.04**	**1.98**	**1.94**	**1.91**
36	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15	2.10	2.06	2.03	1.98	1.93	1.87	1.82	1.78	1.72	1.69	1.65	1.62	1.59	1.56	1.55
	7.39	**5.25**	**4.38**	**3.89**	**3.58**	**3.35**	**3.18**	**3.04**	**2.94**	**2.86**	**2.78**	**2.72**	**2.62**	**2.54**	**2.43**	**2.35**	**2.26**	**2.17**	**2.12**	**2.04**	**2.00**	**1.94**	**1.90**	**1.87**
38	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.96	1.92	1.85	1.80	1.76	1.71	1.67	1.63	1.60	1.57	1.54	1.53
	7.35	**5.21**	**4.34**	**3.86**	**3.54**	**3.32**	**3.15**	**3.02**	**2.91**	**2.82**	**2.75**	**2.69**	**2.59**	**2.51**	**2.40**	**2.32**	**2.22**	**2.14**	**2.08**	**2.00**	**1.97**	**1.90**	**1.86**	**1.84**
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.04	2.00	1.95	1.90	1.84	1.79	1.74	1.69	1.66	1.61	1.59	1.55	1.53	1.51
	7.31	**5.18**	**4.31**	**3.83**	**3.51**	**3.29**	**3.12**	**2.99**	**2.88**	**2.80**	**2.73**	**2.66**	**2.56**	**2.49**	**2.37**	**2.29**	**2.20**	**2.11**	**2.05**	**1.97**	**1.94**	**1.88**	**1.84**	**1.81**
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.02	1.99	1.94	1.89	1.82	1.78	1.73	1.68	1.64	1.60	1.57	1.54	1.51	1.49
	7.27	**5.15**	**4.29**	**3.80**	**3.49**	**3.26**	**3.10**	**2.96**	**2.86**	**2.77**	**2.70**	**2.64**	**2.54**	**2.46**	**2.35**	**2.26**	**2.17**	**2.08**	**2.02**	**1.94**	**1.91**	**1.85**	**1.80**	**1.78**
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.98	1.92	1.88	1.81	1.76	1.72	1.66	1.63	1.58	1.56	1.52	1.50	1.48
	7.24	**5.12**	**4.26**	**3.78**	**3.46**	**3.24**	**3.07**	**2.94**	**2.84**	**2.75**	**2.68**	**2.62**	**2.52**	**2.44**	**2.32**	**2.24**	**2.15**	**2.06**	**2.00**	**1.92**	**1.88**	**1.82**	**1.78**	**1.75**
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.14	2.09	2.04	2.00	1.97	1.91	1.87	1.80	1.75	1.71	1.65	1.62	1.57	1.54	1.51	1.48	1.46
	7.21	**5.10**	**4.24**	**3.76**	**3.44**	**3.22**	**3.05**	**2.92**	**2.82**	**2.73**	**2.66**	**2.60**	**2.50**	**2.42**	**2.30**	**2.22**	**2.13**	**2.04**	**1.98**	**1.90**	**1.86**	**1.80**	**1.76**	**1.72**
48	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08	2.03	1.99	1.96	1.90	1.86	1.79	1.74	1.70	1.64	1.61	1.56	1.53	1.50	1.47	1.45
	7.19	**5.08**	**4.22**	**3.74**	**3.42**	**3.20**	**3.04**	**2.90**	**2.80**	**2.71**	**2.64**	**2.58**	**2.48**	**2.40**	**2.28**	**2.20**	**2.11**	**2.02**	**1.96**	**1.88**	**1.84**	**1.78**	**1.73**	**1.70**

Table A.3 The 5 (Roman Type) and 1 (Boldface Type) Percent Points for the *F* Distribution (*cont.*)

denominator $df = n_1$	n_1 degrees of freedom for numerator																							
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
50	4,03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.02	1.98	1.95	1.90	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.48	1.46	1.44
	7.17	**5.06**	**4.20**	**3.72**	**3.41**	**3.18**	**3.02**	**2.88**	**2.78**	**2.70**	**2.62**	**2.56**	**2.46**	**2.39**	**2.26**	**2.18**	**2.10**	**2.00**	**1.94**	**1.86**	**1.82**	**1.76**	**1.71**	**1.68**
55	4.02	3.17	2.78	2.54	2.38	2.27	2.18	2.11	2.05	2.00	1.97	1.93	1.88	1.83	1.76	1.72	1.67	1.61	1.58	1.52	1.50	1.46	1.43	1.41
	7.12	**5.01**	**4.16**	**3.68**	**3.37**	**3.15**	**2.98**	**2.85**	**2.75**	**2.66**	**2.59**	**2.53**	**2.43**	**2.35**	**2.23**	**2.15**	**2.06**	**1.96**	**1.90**	**1.82**	**1.78**	**1.71**	**1.66**	**1.64**
60	4.00	3.15	2.76	2.52	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.86	1.81	1.75	1.70	1.65	1.59	1.56	1.50	1.48	1.44	1.41	1.39
	7.08	**4.98**	**4.13**	**3.65**	**3.34**	**3.12**	**2.95**	**2.82**	**2.72**	**2.63**	**2.56**	**2.50**	**2.40**	**2.32**	**2.20**	**2.12**	**2.03**	**1.93**	**1.87**	**1.79**	**1.74**	**1.68**	**1.63**	**1.60**
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.02	1.98	1.94	1.90	1.85	1.80	1.73	1.68	1.63	1.57	1.54	1.49	1.46	1.42	1.39	1.37
	7.04	**4.95**	**4.10**	**3.62**	**3.31**	**3.09**	**2.93**	**2.79**	**2.70**	**2.61**	**2.54**	**2.47**	**2.37**	**2.30**	**2.18**	**2.09**	**2.00**	**1.90**	**1.84**	**1.76**	**1.71**	**1.64**	**1.60**	**1.56**
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.01	1.97	1.93	1.89	1.84	1.79	1.72	1.67	1.62	1.56	1.53	1.47	1.45	1.40	1.37	1.35
	7.01	**4.92**	**4.08**	**3.60**	**3.29**	**3.07**	**2.91**	**2.77**	**2.67**	**2.59**	**2.51**	**2.45**	**2.35**	**2.28**	**2.15**	**2.07**	**1.98**	**1.88**	**1.82**	**1.74**	**1.69**	**1.62**	**1.56**	**1.53**
80	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99	1.95	1.91	1.88	1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.42	1.38	1.35	1.32
	6.96	**4.88**	**4.04**	**3.56**	**3.25**	**3.04**	**2.87**	**2.74**	**2.64**	**2.55**	**2.48**	**2.41**	**2.32**	**2.24**	**2.11**	**2.03**	**1.94**	**1.84**	**1.78**	**1.70**	**1.65**	**1.57**	**1.52**	**1.49**
100	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97	1.92	1.88	1.85	1.79	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28
	6.90	**4.82**	**3.98**	**3.51**	**3.20**	**2.99**	**2.82**	**2.69**	**2.59**	**2.51**	**2.43**	**2.36**	**2.26**	**2.19**	**2.06**	**1.98**	**1.89**	**1.79**	**1.73**	**1.64**	**1.59**	**1.51**	**1.46**	**1.43**
125	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.95	1.90	1.86	1.83	1.77	1.72	1.65	1.60	1.55	1.49	1.45	1.39	1.36	1.31	1.27	1.25
	6.84	**4.78**	**3.94**	**3.47**	**3.17**	**2.95**	**2.79**	**2.65**	**2.56**	**2.47**	**2.40**	**2.33**	**2.23**	**2.15**	**2.03**	**1.94**	**1.85**	**1.75**	**1.68**	**1.59**	**1.54**	**1.46**	**1.40**	**1.37**
150	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82	1.76	1.71	1.64	1.59	1.54	1.47	1.44	1.37	1.34	1.29	1.25	1.22
	6.81	**4.75**	**3.91**	**3.44**	**3.14**	**2.92**	**2.76**	**2.62**	**2.53**	**2.44**	**2.37**	**2.30**	**2.20**	**2.12**	**2.00**	**1.91**	**1.83**	**1.72**	**1.66**	**1.56**	**1.51**	**1.43**	**1.37**	**1.33**
200	3.89	3.04	2.65	2.41	2.26	2.14	2.05	1.98	1.92	1.87	1.83	1.80	1.74	1.69	1.62	1.57	1.52	1.45	1.42	1.35	1.32	1.26	1.22	1.19
	6.76	**4.71**	**3.88**	**3.41**	**3.11**	**2.90**	**2.73**	**2.60**	**2.50**	**2.41**	**2.34**	**2.28**	**2.17**	**2.09**	**1.97**	**1.88**	**1.79**	**1.69**	**1.62**	**1.53**	**1.48**	**1.39**	**1.33**	**1.28**
400	3.86	3.02	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.78	1.72	1.67	1.60	1.54	1.49	1.42	1.38	1.32	1.28	1.22	1.16	1.13
	6.70	**4.66**	**3.83**	**3.36**	**3.06**	**2.85**	**2.69**	**2.55**	**2.46**	**2.37**	**2.29**	**2.23**	**2.12**	**2.04**	**1.92**	**1.84**	**1.74**	**1.64**	**1.57**	**1.47**	**1.42**	**1.32**	**1.24**	**1.19**
1000	3.85	3.00	2.61	2.38	2.22	2.10	2.02	1.95	1.89	1.84	1.80	1.76	1.70	1.65	1.58	1.53	1.47	1.41	1.36	1.30	1.26	1.19	1.13	1.08
	6.66	**4.62**	**3.80**	**3.34**	**3.04**	**2.82**	**2.66**	**2.53**	**2.43**	**2.34**	**2.26**	**2.20**	**2.09**	**2.01**	**1.89**	**1.81**	**1.71**	**1.61**	**1.54**	**1.44**	**1.38**	**1.28**	**1.19**	**1.11**
∞	3.84	2.99	2.60	2.37	2.21	2.09	2.01	1.94	1.88	1.83	1.79	1.75	1.69	1.64	1.57	1.52	1.46	1.40	1.35	1.28	1.24	1.17	1.11	1.00
	6.64	**4.60**	**3.78**	**3.32**	**3.02**	**2.80**	**2.64**	**2.51**	**2.41**	**2.32**	**2.24**	**2.18**	**2:07**	**1.99**	**1.87**	**1.79**	**1.69**	**1.59**	**1.52**	**1.41**	**1.36**	**1.25**	**1.15**	**1.00**

Source: Reprinted by permission from *Statistical Methods*, 8th ed., by George W. Snedecor and William G. Cochran. Copyright © 1980 by Iowa State University Press, Ames, Iowa 50010.

Table A.4 Table of χ^2

df	.99	.98	.95	.90	.80	.70	.50	.30	.20	.10	.05	.02	.01	.001
1	.000157	.000628	.00393	.0158	.0642	.148	.455	1.074	1.642	2.706	3.841	5.412	6.635	10.827
2	.0201	.0404	.103	.211	.446	.713	1.386	2.408	3.219	4.605	5.991	7.824	9.210	13.815
3	.115	.185	.352	.584	1.005	1.424	2.366	3.665	4.642	6.251	7.815	9.837	11.345	16.266
4	.297	.429	.711	1.064	1.649	2.195	3.357	4.878	5.989	7.779	9.488	11.668	13.277	18.467
5	.554	.752	1.145	1.610	2.343	3.000	4.351	6.064	7.289	9.236	11.070	13.388	15.086	20.515
6	.872	1.134	1.635	2.204	3.070	3.828	5.348	7.231	8.558	10.645	12.592	15.033	16.812	22.457
7	1.239	1.564	2.167	2.833	3.822	4.671	6.346	8.383	9.803	12.017	14.067	16.622	18.475	24.322
8	1.646	2.032	2.733	3.490	4.594	5.527	7.344	9.524	11.030	13.362	15.507	18.168	20.090	26.125
9	2.088	2.532	3.325	4.168	5.380	6.393	8.343	10.656	12.242	14.684	16.919	19.679	21.666	27.877
10	2.558	3.059	3.940	4.865	6.179	7.267	9.342	11.781	13.442	15.987	18.307	21.161	23.209	29.588
11	3.053	3.609	4.575	5.578	6.989	8.148	10.341	12.899	14.631	17.275	19.675	22.618	24.725	31.264
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340	14.011	15.812	18.549	21.026	24.054	26.217	32.909
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340	15.119	16.985	19.812	22.362	25.472	27.688	34.528
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339	16.222	18.151	21.064	23.685	26.873	29.141	36.123
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339	17.322	19.311	22.307	24.996	28.259	30.578	37.697

Table A.4 **Table of χ^2 (*cont.*)**

df	.99	.98	.95	.90	.80	.70	.50	.30	.20	.10	.05	.02	.01	.001
16	5.812	6.614	7.962	9.312	11.152	12.624	15.338	18.418	20.465	23.542	26.296	29.633	32.000	39.252
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338	19.511	21.615	24.769	27.587	30.995	33.409	40.790
18	7.015	7.906	9.390	10.865	12.857	14.440	17.338	20.601	22.760	25.989	28.869	32.346	34.805	42.312
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338	21.689	23.900	27.204	30.144	33.687	36.191	43.820
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337	22.775	25.038	28.412	31.410	35.020	37.566	45.315
21	8.897	9.915	11.591	13.240	15.445	17.182	20.337	23.858	26.171	29.615	32.671	36.343	38.932	46.797
22	9.542	10.600	12.338	14.041	16.314	18.101	21.337	24.939	27.301	30.813	33.924	37.659	40.289	48.268
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337	26.018	28.429	32.007	35.172	38.968	41.638	49.728
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337	27.096	29.553	33.196	36.415	40.270	42.980	51.179
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337	28.172	30.675	34.382	37.652	41.566	44.314	52.620
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336	29.246	31.795	35.563	38.885	42.856	45.642	54.052
27	12.879	14.125	16.151	18.114	20.703	22.719	26.336	30.319	32.912	36.741	40.113	44.140	46.963	55.476
28	13.565	14.847	16.928	18.939	21.588	23.647	27.336	31.391	34.027	37.916	41.337	45.419	43.278	56.893
29	14.256	15.574	17.708	19.768	22.475	24.577	28.336	32.461	35.139	39.087	42.557	46.693	49.588	58.302
30	14.953	16.306	18.493	20.599	23.364	25.508	29.336	33.530	36.250	40.256	43.773	47.962	50.892	59.703

Source: Taken from Table IV of Fisher and Yates: *Statistical Tables for Biological, Agricultural and Medical Research,* published by Longman Group UK Ltd., 1974.

Table A.5 Critical Values of the Pearson Product Moment Correlation Coefficient

	Level of significance for a directional (one-tailed) test				
	.05	.025	.01	.005	.0005
	Level of significance for a nondirectional (two-tailed) test				
$df = N - 2$	.10	.05	.02	.01	.001
1	.9877	.9969	.9995	.9999	1.0000
2	.9000	.9500	.9800	.9900	.9990
3	.8054	.8783	.9343	.9587	.9912
4	.7293	.8114	.8822	.9172	.9741
5	.6694	.7545	.8329	.8745	.9507
6	.6215	.7067	.7887	.8343	.9249
7	.5822	.6664	.7498	.7977	.8982
8	.5494	.6319	.7155	.7646	.8721
9	.5214	.6021	.6851	.7348	.8471
10	.4973	.5760	.6581	.7079	.8233
11	.4762	.5529	.6339	.6835	.8010
12	.4575	.5324	.6120	.6614	.7800
13	.4409	.5139	.5923	.6411	.7603
14	.4259	.4973	.5742	.6226	.7420
15	.4124	.4821	.5577	.6055	.7246
16	.4000	.4683	.5425	.5897	.7084
17	.3887	.4555	.5285	.5751	.6932
18	.3783	.4438	.5155	.5614	.6787
19	.3687	.4329	.5034	.5487	.6652
20	.3598	.4227	.4921	.5368	.6524
25	.3233	.3809	.4451	.4869	.5974
30	.2960	.3494	.4093	.4487	.5541
35	.2746	.3246	.3810	.4182	.5189
40	.2573	.3044	.3578	.3932	.4896
45	.2428	.2875	.3384	.3721	.4648
50	.2306	.2732	.3218	.3541	.4433
60	.2108	.2500	.2948	.3248	.4078
70	.1954	.2319	.2737	.3017	.3799
80	.1829	.2172	.2565	.2830	.3568
90	.1726	.2050	.2422	.2673	.3375
100	.1638	.1946	.2301	.2540	.3211

Source: Table A.5 is taken from Table VII of Fisher and Yates, *Statistical Tables for Biological, Agricultural, and Medical Research*, published by Longman Group UK Ltd., 1974.

Table A.6 Table of Random Numbers

Row	00000 01234	00000 56789	11111 01234	11111 56789	22222 01234	22222 56789	33333 01234	33333 56789
				1st Thousand				
00	23157	54859	01837	25993	76249	70886	95230	36744
01	05545	55043	10537	43508	90611	83744	10962	21343
02	14871	60350	32404	36223	50051	00322	11543	80834
03	38976	74951	94051	75853	78805	90194	32428	71695
04	97312	61718	99755	30870	94251	25841	54882	10513
05	11742	69381	44339	30872	32797	33118	22647	06850
06	43361	28859	11016	45623	93009	00499	43640	74036
07	93806	20478	38268	04491	55751	18932	58475	52571
08	49540	13181	08429	84187	69538	29661	77738	09527
09	36768	72633	37948	21569	41959	68670	45274	83880
10	07092	52392	24627	12067	06558	45344	67338	45320
11	43310	01081	44863	80307	52555	16148	89742	94647
12	61570	06360	06173	63775	63148	95123	35017	46993
13	31352	83799	10779	18941	31579	76448	62584	86919
14	57048	86526	27795	93692	90529	56546	35065	32254
15	09243	44200	68721	07137	30729	75756	09298	27650
16	97957	35018	40894	88329	52230	82521	22532	61587
17	93732	59570	43781	98885	56671	66826	95996	44569
18	72621	11225	00922	68264	35666	59434	71687	58167
19	61020	74418	45371	20794	95917	37866	99536	19378
20	97839	85474	33055	91718	45473	54144	22034	23000
21	89160	97192	22232	90637	35055	45489	88438	16361
22	25966	88220	62871	79265	02823	52862	84919	54883
23	81443	31719	05049	54806	74690	07567	65017	16543
24	11322	54931	42362	34386	08624	97687	46245	23245

Table A.6 Table of Random Numbers (*cont.*)

Row	Column Number 00000 01234	00000 56789	11111 01234	11111 56789	22222 01234	22222 56789	33333 01234	33333 56789
				2nd Thousand				
00	64755	83885	84122	25920	17696	15655	95045	95947
01	10302	52289	77436	34430	38112	49067	07348	23328
02	71017	98495	51308	50374	66591	02887	53765	69149
03	60012	55605	88410	34879	79655	90169	78800	03666
04	37330	94656	49161	42802	48274	54755	44553	65090
05	47869	87001	31591	12273	60626	12822	34691	61212
06	38040	42737	64167	89578	39323	49324	88434	38706
07	73508	30908	83054	80078	86669	30295	56460	45336
08	32623	46474	84061	04324	20628	37319	32356	43969
09	97591	99549	36630	35106	62069	92975	95320	57734
10	74012	31955	59790	96982	66224	24015	96749	07589
11	56754	26457	13351	05014	90966	33674	69096	33488
12	49800	49908	54831	21998	08528	26372	92923	65026
13	43584	89647	24878	56670	00221	50193	99591	62377
14	16653	79664	60325	71301	35742	83636	73058	87229
15	48502	69055	65322	58748	31446	80237	31252	96367
16	96765	54692	36316	86230	48296	38352	23816	64094
17	38923	61550	80357	81784	23444	12463	33992	28128
18	77958	81694	25225	05587	51073	01070	60218	61961
19	17928	28065	25586	08771	02641	85064	65796	48170
20	94036	85978	02318	04499	41054	10531	87431	21596
21	47460	60479	56230	48417	14372	85167	27558	00368
22	47856	56088	51992	82439	40644	17170	13463	18288
23	57616	34653	92298	62018	10375	76515	62986	90756
24	08300	92704	66752	66610	57188	79107	54222	22013

Table A.6 Table of Random Numbers (*cont.*)

Row	Column Number 00000 01234	00000 56789	11111 01234	11111 56789	22222 01234	22222 56789	33333 01234	33333 56789
				3rd Thousand				
00	89221	02362	65787	74733	51272	30213	92441	39651
01	04005	99818	63918	29032	94012	42363	01261	10650
02	98546	38066	50856	75045	40645	22841	53254	44125
03	41719	84401	59226	01314	54581	40398	49988	65579
04	28733	72489	00785	25843	24613	49797	85567	84471
05	65213	83927	77762	03086	80742	24395	68476	83792
06	65553	12678	90906	90466	43670	26217	69900	31205
07	05668	69080	73029	85746	58332	78231	45986	92998
08	39202	99718	49757	79519	27387	76373	47262	91612
09	64592	32254	45879	29431	38320	05981	18067	87137
10	07513	48792	47314	83660	68907	05336	82579	91582
11	86593	68501	56638	99800	82839	35148	56541	07232
12	83735	22599	97977	81248	36838	99560	32410	67614
13	08595	21826	54655	08204	87990	17033	56258	05384
14	41273	27149	44293	69458	16828	63962	15864	35431
15	00473	75908	56238	12242	72631	76314	47252	06347
16	86131	53789	81383	07868	89132	96182	07009	86432
17	33849	78359	08402	03586	03176	88663	08018	22546
18	61870	41657	07468	08612	98083	97349	20775	45091
19	43898	65923	25078	86129	78491	97653	91500	80786
20	29939	39123	04548	45985	60952	06641	28726	46473
21	38505	85555	14388	55077	18657	94887	67831	70819
22	31824	38431	67125	25511	72044	11562	53279	82268
23	91430	03767	13561	15597	06750	92552	02391	38753
24	38635	68976	25498	97526	96458	03805	04116	63514

Table A.6 Table of Random Numbers (*cont.*)

Row	Column Number 00000 01234	00000 56789	11111 01234	11111 56789	22222 01234	22222 56789	33333 01234	33333 56789
				4th Thousand				
00	02490	54122	27944	39364	94239	72074	11679	54082
01	11967	36469	60627	83701	09253	30208	01385	37482
02	48256	83465	49699	24079	05403	35154	39613	03136
03	27246	73080	21481	23536	04881	89977	49484	93071
04	32532	77265	72430	70722	86529	18457	92657	10011
05	66757	98955	92375	93431	43204	55825	45443	69265
06	11266	34545	76505	97746	34668	26999	26742	97516
07	17872	39142	45561	80146	93137	48924	64257	59284
08	62561	30365	03408	14754	51798	08133	61010	97730
09	62796	30779	35497	70501	30105	08133	00997	91970
10	75510	21771	04339	33660	42757	62223	87565	48468
11	87439	01691	63517	26590	44437	07217	98706	39032
12	97742	02621	10748	78803	38337	65226	92149	59051
13	98811	06001	21571	02875	21828	83912	85188	61624
14	51264	01852	64607	92553	29004	26695	78583	62998
15	40239	93376	10419	68610	49120	02941	80035	99317
16	26936	59186	51667	27645	46329	44681	94190	66647
17	88502	11716	98299	40974	42394	62200	69094	81646
18	63499	38093	25593	61995	79867	80569	01023	38374
19	36379	81206	03317	78710	73828	31083	60509	44091
20	93801	22322	47479	57017	59334	30647	43061	26660
21	29856	87120	56311	50053	25365	81265	22414	02431
22	97720	87931	88265	13050	71017	15177	06957	92919
23	85237	09105	74601	46377	59938	15647	34177	92753
24	75746	75268	31727	95773	72364	87324	36879	06802

Table A.6 Table of Random Numbers (*cont.*)

Row	Column Number 00000 01234	00000 56789	11111 01234	11111 56789	22222 01234	22222 56789	33333 01234	33333 56789
				5th Thousand				
00	29935	06971	63175	52579	10478	89379	61428	21363
01	15114	07126	51890	77787	75510	13103	42942	48111
02	03870	43225	10589	87629	22039	94124	38127	65022
03	79390	39188	40756	45269	65959	20640	14284	22960
04	30035	06915	79196	54428	64819	52314	48721	81594
05	29039	99861	28759	79802	68531	39198	38137	24373
06	78196	08108	24107	49777	09599	43569	84820	94956
07	15847	85493	91442	91351	80130	73752	21539	10986
08	36614	62248	49194	97209	92587	92053	41021	80064
09	40549	54884	91465	43862	35541	44466	88894	74180
10	40878	08997	14286	09982	90308	78007	51587	16658
11	10229	49282	41173	31468	59455	18756	08908	06660
12	15918	76787	30624	25928	44124	25088	31137	71614
13	13403	18796	49909	94404	64979	41462	18155	98335
14	66523	94596	74908	90271	10009	98648	17640	68909
15	91665	36469	68343	17870	25975	04662	21272	50620
16	67415	87515	08207	73729	73201	57593	96917	69699
17	76527	96996	23724	33448	63392	32394	60887	90617
18	19815	47789	74348	17147	10954	34355	81194	54407
19	25592	53587	76384	72575	84347	68918	05739	57222
20	55902	45539	63646	31609	95999	82887	40666	66692
21	02470	58376	79794	22482	42423	96162	47491	17264
22	18630	53263	13319	97619	35859	12350	14632	87659
23	89673	38230	16063	92007	59503	38402	76450	33333
24	62986	67364	06595	17427	84623	14565	82860	57300

Source: Kendall, M. G., and B. B. Smith (1938). Randomness and random sampling numbers. *Journal of the Royal Statistical Society, 101*, 164–166. Reproduced by permission of the Royal Statistical Society, London.

المسرد (قائمة شرح المصطلحات) GLOSSARY

- A -

تصميم ABAB **ABAB design**

تصميم تجريبي للفرد الواحد حيث تتم قياسات خط القاعدة للسلوك المستهدف (A) يليها المعالجة (B)، ثم قياس ثان بعدئذ لخط القاعدة (A) وقياس ثان للمعالجة (B).

خلاصة (وجيزة) **Abstract**

خلاصة مختصرة لمحتويات وثيقة

مجتمع سهل المنال / الوصول إليه **Accessible population**

مجتمع الأفراد الذي يتيسّر للباحث الوصول إليه لدراسة ما، وهو المجتمع الذي يمكن تعميم النتائج عليه.

معاينة عرضية/ تصادفية **Accidental sampling**

أسلوب معاينة لا احتمالي يستخدم الأفراد المتيسرين دون اعتبار لدى تمثيل العينة للمجتمع المعني

اختبار تحصيل **Achievement test**

مقياس لمدى ما اكتسبه الفرد من معلومات، ومهارات، تأتي غالبا عن تدريس محدد.

متغير مستقل فعّال/ نشيط **Active independen variable**

متغير مستقل ناجم خلال التأثير النشط في مسار التجربة. والمتغيرات المستقلة الفعّالة خاصة للبحث التجريبي

معامل الاتفاق **Agreement coefficient**

نسبة الأفراد الذين يتخذ القرار ذاته بشأنهم (إتقان أو لا إتقان) في إجرائين لاختبار مرجعي الإتقان (المحك). ويستخدم كمؤشر لثبات هذه الاختبارات.

معامل ألفا **Alpha coefficient**

أنظر Cronbach alpha

الصورة البديلة **Alternate-form**

أنظر Equivalent-form

تحليل التباين الاقتراني **Analysis of covariance (ANOVA)**

أسلوب إحصائي يوفر ضبطا إحصائياً جزئياً لمتغير أو أكثر، بإقصاء تأثيرها من مقارنة المجموعات حسب المتغير التابع.

تحليل التباين **Analysis of variance**

اختبار استدلال إحصائي يُستخدم للتصميمات التجريبية التي لها أكثر من متغير مستقل أو أكثر من مستويين لمتغير مستقل.

بحث تطبيقي **Applied research**

بحث يهدف إلى حل مشكلة عملية عاجلة.

اختبار استعداد **Aptitude test**

اختبار يقيس قدرات عامة أو خصائص يُعتقد أنها دالة على قدرة شخص لتعلم مطلب مستقبلي أو لإنجاز في مجال معين.

مقياس اتجاهات **Attitude scale**

مقياس درجة محاباة/ تأييد أو عدم محاباة شخص نحو مجموعة، مؤسسة، مفهوم بنائي، أو شئ ما.

متغير مستقل منسوب **Attribute independent variable**

متغير مستقل يختلف فيه الأفراد قبل بداية الدراسة. وهذه المتغيرات خصائص للبحث العلّي – المقارن.

- B -

خط القاعدة/ الأساس **Baseline**

قياسات للمتغير التابع تؤخذ قبل تقديم المعالجة في تصميم سلسلة زمنية تجريبي وتستخدم كمعيار للمقارنة.

بحث أساسي **Basic research**

بحث يهدف إلى الحصول على بيانات تجريبية (empirical) تستخدم لصياغة، توسيع، أو تقويم نظرية وليس حل مشكلة عملية.

عينة متحيزة/ مغرضة **Based sample**

عينة يكون فيها تمثيل عناصر معينة في المجتمع الأصلي دون أو فوق مستوى تمثيلها.

ارتباط ثنائي **Biserial correlation**

إجراء إحصائي لإيجاد العلاقة بين متغير متصل/ مستمر ومتغير آخر منفصل جرى تصنيفه/ تقسيمه ثنائيا على نحو مصطنع.

المنطق البولي **Boolian logic**

نظام يستخدم في استقصاءات الحاسوب للأدبيات لتحديد وثائق عن طريق كلمات رئيسية مفتاحية خاصة.

- C -

ارتباط ممتد/ موسع **Canonical correlation**

إجراء إحصائي لتحديد العلاقة بين عدة متغيرات مستقلة وأكثر من متغير تابع.

دراسة الحالة **Case study**

استقصاء نوعي لفرد واحد أو مجموعة

متغيرات نوعية/ طبقية **Categorical variables**

متغيرات تختلف في النوع/ الصنف، وليس في الدرجة أو المقدار

البحث العلّي — المقارن **Causal-comparative research**

نوع من البحث يسعى لتحديد العلل، أو العواقب/ النتائج، لفروق توجد مسبقاً في مجموعات الأفراد. ويدعي أيضاً: البحث اللاحق للحادث (Ex post facto research)

علاقة علّية **Causal relationship**

علاقة ينشأ فيها، عن التغيرات في متغير واحد، تغيرات في متغير آخر.

إحصاء السكان **Census**

مسح يشمل كل المجتمع المعنى.

النزعة المركزية (قياس) **Central tendency (Measure of)**

تستخدم المتوسطات مثل الوسط (الحسابي)، والوسيط، والمنوال لتلخيص البيانات في توزيع تكراري.

درجة التغير **Change score**

الفرق بين درجات الأفراد في الاختبارين القبلي والبعدي للمتغير التابع

Checklist **قائمة تدقيق/ مراجعة**

نمط من سؤال مسح يعرض أعداداً للإجابات الممكنة، ويطلب إلى المستجيبين تحديد ما ينطبق.

Chi-square (X^2) **مربع كاي**

إحصاءة استدلالية تقارن التكرارات/ القياسات الاسمية التي تمت ملاحظتها فعلاً في دراسة ما مع التكرارات المتوقعة تحت فرضية صفرية (صدفة).

CIJE (Current Index to Journal in Education) **الفهرس/ الدليل الراهن للمجلات في التربية**

مجلة ERIC الشهرية التي تفهرس مقالات من بين أكثر من 700 مجلة حسب الموضوع، المؤلف، والمجلة، وتحوى خلاصات وجيزة لهذه المقالات.

Closed - ended question **سؤال مغلق النهاية**

سؤال تتبعه مجموعة محددة من الاستجابات البديلة ليختار منها المستجيب.

Cluster sampling **معاينة عنقودية**

أسلوب معاينة احتمالية يختار عشوائياً مجموعات تامة (عناقيد) كصفوف مدرسية أو دوائر انتخابية ثم يشمل كل عنصر في العنقود المختار في العينة.

Coefficient of correlation **معامل الارتباط**

أنظر Correlation coefficient

Coefficient of determinism **معامل التحديد**

مربع معامل الارتباط. ويشير إلى النسبة المئوية للتباين في متغير يشترك مع متغير آخر.

Cohort study **دراسة الجماعة**

دراسة مسحية طولية تتم بدراسة مجتمع خاص/ محدد بأخذ عينات عشوائية مختلفة من المجتمع في نقاط زمنية متباينة.

Common cause **علّة مشتركة**

متغير يتسبب في فروقات متغيرين مفروضين، مستقل وتابع.

Comparison group **مجموعة مقارنة**

المجموعة في دراسة التي لا تتلقى أية معالجة، أو تتلقى معالجة مختلفة عن المعالجة التي تتلقاها المجموعة التجريبية.

بحث حاسوبي **Computer search**

استخدام حاسوب لتحديد معلومات في قاعدة بيانات.

مفهوم **Concept**

تجريد يمثل مجموعة من الأشياء أو الأحداث التي تشترك في شئ ما.

صدق تلازمي **Concurrent validity**

أنظر Criterion-related validity

متغير مربك/ مشوش **Confounding variable**

متغير دخيل غير منضبط قد تُعزي، خطأً، تأثيراته على المتغير التابع للمتغير المستقل.

ثابت **Constant**

خاصة تتخذ القيمة ذاتها لجميع الأفراد في دراسة ما، خلافاً للمتغير.

تعريف تأسيسي (قاموسي) **Constitutive definition**

تعريف يتم فيه تعريف كلمة باستخدام كلمات أخرى

مفهوم بنائي **Construct**

تجريد أعلى مستوى من المفهوم (concept) يستخدم لتوضيح، وتفسير، وتلخيص ملاحظات ليشكل جزءاً من المحتوى القائم على مفاهيم لنظرية ما.

(دليل) الصدق المتعلق بالمفهوم البنائي **Construct-related validity (evidence of)**

المدى الذي يمكن لأداة أن تقيس سمة أو خاصة ينطوي عليها المفهوم البنائي المقصود قياسه.

تحليل المحتوي **Content analysis**

أنظر: Documentary analysis

(دليل) الصدق المتعلق بالمحتوى/ المضمون **Content-related validity (Evidence of)**

المدى الذي تمثل فيه بنود أداة عينة، مجال المحتوى الأساسي.

سؤال موقوف/ مُعلّق **Contingency question**

سؤال مسحي تتوقف/ تتعلق الإجابة عنه على إجابة لسؤال سابق

متغير متصل/ مستمر **Continuous variable**

متغير يمكن أن يأخذ قياسه عدداً غير منته من النقاط (القيم) في مدى معين.

ملاحظة مخططة **Contrived observation**

ملاحظة/ رصد لأفراد في موقف مرتب/ مخطط مسبقاً.

ضبط/ تحكم **Control**

خطوات يتخذها الباحث لاستبعاد تأثير متغير (أو أكثر) عدا المتغير المستقل الذي قد يؤثر في المتغير التابع.

مجموعة ضابطة **Control group**

هل المجموعة التي لا تتلقى، في دراسة، المعالجة التجريبية، وتتم مقارنتها بالمجموعة التجريبية لتحديد تأثيرات المعالجة.

الارتباط/ الترابط **Correlation**

أسلوب لتحديد التغيرات المترافقة/ المقترنة بين مجموعات من الدرجات، فأزواج الدرجات قد تتغير طردا (زيادة أو نقصان معاً) أو تتغير عكسياً (عندما تزيد واحدة تنقص الأخرى).

البحث الارتباطي **Correlational research**

بحث يسعى لتحديد درجة/ مدى واتجاه العلاقة بين متغيرين أو أكثر

معامل ارتباط **Correlation coefficient**

إحصاءة (statistic) تبين درجة العلاقة بين متغيرين، وتتراوح قيمتها بين 1.00- و 1.00+

مصفوفة الارتباط **Correlation matrix**

جدول يوضح معاملات الارتباط بين كل مقياس وآخر

تصميم الموازنة/ التدوير **Counterbalanced design**

تصميم تجريبي يتم فيه ضبط تأثيرات الترتيب بجعل كل المجموعات تتلقى كل المعالجات حسب ترتيب مختلف.

اختبار مرجعي الإتقان **Criterion-referenced test**

أداة تقيس مستوى الإتقان في مجال محتوى أو مهارة معرفتين مسبقاً.

(هناك من يترجم هذا المصطلح "اختبار مرجعي المحك" والمحك هنا هو معيار يحدد الإتقان، لكن المحك/ المعيار كلمتان غير محددتين بما يثير الغموض لدى القارئ – المراجع).

(دليل) الصدق المتعلق بالمعيار **Criterion-related validity (evidence of)**

هو مدى ارتباط أداة بمؤشرات للشيء ذاته (المعيار). فإذا تم جمع درجات المعيار في الوقت ذاته، فإن الارتباط دليل "للصدق التلازمي" أما إذا جمعت الدرجات في وقت لاحق، فإن الارتباط الناجم بين الدرجات والمعيار يكون دليلاً "للصدق التنبؤي".

معامل الارتباط ألفا (كرونباخ) **Cronbach alpha**

معامل ثبات للتوافق الداخلي، يقيس مدى اتفاق درجات الفرد مع بعضها. وهو مفيد بوجه خاص في مقاييس الاتجاهات واختبارات المقالة.

مسح (مقطع) عرضي **Cross-sectional survey**

مسح يتم فيه جمع البيانات في نقطة واحدة زمنية من مجتمع محدد.

جدولة متقاطعة **Cross-tabulation**

جدول يوضح مدى تكرار حدوث جميعة/ توليفة من متغيرين أو أكثر، حيث يمكن للمرء أن "يدرك" منها العلاقة (إن وجدت) بين المتغيرات.

تحقيق متقاطع / تبادلي للصدق **Cross-validation**

عملية تحقيق معاجلة التنبؤ القائمة على مجموعة باستخدامها مع مجموعة مختلفة.

تكرار تراكمي **Cumulative frequency**

عدد الحالات التي تكون درجاتها، في توزيع ما، مساوية أو أقل من درجة محددة.

علاقة منحنية **Curvilinear relationship**

علاقة ارتباط، حيث يوصف بيان/ مبيان التشتت لأزواج الدرجات على نحو أفضل بمنحن وليس بمستقيم. فمثلاً، إذا جرت مزاوجة الدرجات العالية في Y مع الدرجات المتوسطة في X، والدرجات المتدنية في Y مع الدرجات العالية والمتدنية في X، فإن العلاقة تكون محنية.

- D -

استنتاج/ استنباط **Deduction**

البدء من مقدمات عامة أو حقائق معروفة مسبقاً واشتقاق نتائج منطقية محددة.

فرضية استنتاجية **Deductive hypothesis**

فرضية تشتق بالاستنتاج من نظرية

درجات حرية **Degrees of freedom (df)**

عدد الملاحظات التي لها حرية التغير حول معْلم/ بارامتر ثابت. ولكل إحصاءة استدلالية إجراء محدد لحساب درجات الحرية لها، حيث يتم استخدامها لتحديد القيم الحرجة المناسبة في جداول إحصائية لتقييم الإحصاءة.

متغير تابع **Dependent variable**

متغير يكون ناتجاً عن، أو معتمداً على، متغير (مستقل) سابق له، ويدعى أيضاً "الناتج" أو "متغير الأثر/ الناتج"

بحث وصفي **Descriptive research**

بحث يطرح أسئلة حول طبيعة، أو وقوع، أو توزيع المتغيرات، ويتضمن وصفاً وليس تفعيلاً للمتغيرات.

إحصاء وصفي **Descriptive statistics**

أساليب لتنظيم وتلخيص ووصف الملاحظات.

دراسة النمو/ التطور **Developmental study**

استقصاء حول كيفية تغير الأطفال حتى مرحلة النضج في استجاباتهم لبيئتهم.

متغير ثنائي **Dichotomous variable**

متغير نوعي/ طبقي له صنفان فقط.

اختيار فارقي/ تمييزي **Differential selection**

تحديد مجموعتين تجريبية وضابطة بطريقة تكونان فيها مختلفتين قبل المعالجة، وهي مشكلة صدق داخلي.

Directional hypothesis

فرضية تحدد الاتجاه المتوقع للاستنتاجات – أي، فيما إذا كانت نتيجة متوقعة "أكبر من" أو "أصغر من".

اختبار أحادي الطرف/ الذيل **Directional test (one-tailed test)**

اختبار إحصائي لفرضية تحدد، فقط أن معْلم/ بارامتر المجتمع الإحصائي أكبر أو أصغر من قيمة المعْلم الذي تحدده الفرضية الصفرية.

ملاحظة مباشرة **Direct observation**

ملاحظة / رصد الأفراد بهدف تسجيل تكرار وزمن استمرار أنواع السلوك (السلوكيات) لتي تقع في أصناف محددة مسبقاً.

التحليل المميِّز/ التمييزي **Discriminant analysis**

إجراء ارتباطي يستخدم عدداً من عوامل / متغيرات التنبؤ بغية تنبؤ العضوية في متغيرات تابعة نوعية.

التحليل الوثائقي **Documentary analysis**

الفحص المنظم/ المنهجي للوثائق بغية استقصاء المواضيع أو الأفكار.

سؤال مزدوج **Double barreled question**

سؤال مسح واحد يطرح فعلاً سؤالين في واحد.

تجربة ثنائية التعمية **Double-blind experiment**

تجربة يجهل فيه الملاحظون والأفراد معرفة أي من المجموعتين التجريبية والضابطة.

متغير مصطنع / صوري **Dummy variable**

متغير نوعي جرى تسجيله كمتغير ثنائي/ مزدوج للدخول في الارتداد المتعدد. فيمكن مثلاً، ترميز الجنس في 1 أو صفر.

- E -

صدق بيئي/ تبيّـؤ **Ecological validity**

هو المدى الذي تكون فيه، نتائج تجربة، مستقلة عن محيط/ خلفية بحث، وبالوسع تعميمها إلى أوضاع أخرى.

حجم الأثر/ التأثير **Effect size**

خارج قسمة الفرق بين المجموعتين التجريبية والضابطة على الانحراف المعياري للمجموعة الضابطة.

تصميم طارئ/ متدرج **Emergent design**

بحث قد يتغير فيه تصميم الاستقصاء مع سير البحث.

صورة مكافئة **Equivalent-form**

إجراء تقييمي للثبات عن طريق ارتباط درجات الأفراد أنفسهم في اختبارين متماثلين، قدر الإمكان، في المحتوى، والصعوبة، والطول، والبنية، وغيرها.

ERIC (Educational Resources Information Center)

إرك (مركز مصادر المعلومات التربوية)

وكالة المكتب الأمريكي للتربية يقوم بجمع المعلومات التربوية، وحفظها وتنظيمها كما يتيح توفير هذه المعلومات للجمهور.

خطأ النـــزعة المركزية **Error of central tendency**

نزعة المقدّر لتجنب المواقف المتطرفة في المقياس حيث يكون تقدير الأفراد قريباً من وسط المقياس.

خطأ التشدّد/ الصرامة **Error of severity**

نزعة لتقدير كل الأفراد على نحو متدنٍ جداً.

أثنوجرافيا/ علم وصف الأعراق **Ethnography**

أحد أشكال البحث النوعي الذي يهدف إلى تكوين صورة شاملة عن مجموعات ثقافية، باستخدام مقابلات معمقة وملاحظة مشارك مطولة.

تصميم تجريبي **Experimental design**

خطة لتجربة تحدد المتغيرات المستقلة التي سيجري تطبيقها، وعدد مستويات كل منها، وكيفية تخصيص الأفراد في المجموعات، كما تحدد المتغير التابع.

مجموعة تجريبية **Experimental group**

المجموعة، في دراسة بحثية، التي تتلقى المعالجة التجريبية.

تسرب تجريبي **Experimental mortality**

تآكل/ فقدان الأفراد خلال دراسة ما. وهذا يتسبب في مشكلة صدق داخلي إذا لم يكن التآكل متماثلاً في المجموعتين التجريبية والضابطة.

تعريف إجرائي تجريبي **Experimental operation definition**

تعريف يصاغ بدلالة الخطوات المتخذة لإنتاج ظروف تجريبية معينة.

بحث تجريبي **Experimental research**

بحث يقوم فيه الباحث بتفعيل واحد أو أكثر من متغيرات مستقلة (المعالجة)، ويلاحظ الأثر على واحد أو أكثر من متغيرات تابعة.

تحيز تجريبي **Experimental bias**

تأثيرات اتجاهات القائم بالتجربة وسلوكه وتوقعاته على سلوك الأفراد في التجربة.

Ex post facto research

راجع Sausal-conparative research

نقد خارجي **External criticism**

تقييم أصالة الوثيقة أو أي دليل مادي آخر في البحث التاريخي

صدق خارجي **External validity**

المدى الذي يسمح بتعميم دراسة خاصة على أفراد آخرين، أوضاع أخرى، و/ أو تعريفات إجرائية أخرى للمتغيرات.

صدق خارجي للإجراءات **External validity of operations**

المدى الذي تكون فيه، التعريفات الإجرائية والطريقة التي تم تمت فيها، تمثل فعلاً متغيرات الدراسة.

متغير دخيل/ خارجي **Extraneous variable**

متغير، لا يخضع للضبط، قد يؤثر على المتغير التابع في الدراسة، وقد يُعزى تأثيره خطأً للمتغير المستقل في الدراسة.

- F -

عامل **Factor**

مفهوم بنائي أساسي، يعلّل التغاير بين عدد أكبر من المتغيرات. ويستخدم أيضاً كمرادف للمتغير المستقل.

تحليل عاملي **Factor analysis**

أسلوب إحصائي لتحليل الارتباطات الداخلية بين ثلاثة مقاييس أو أكثر بما يخفض المجموعة إلى عدد أصغر من العوامل الأساسية.

تصميم عاملي **Factorial design**

تصميم تجريبي يستقصي متغيرين مستقلين أو أكثر في الوقت ذاته لكي يدرس تأثير كل منها على انفراد، و/ أو في التفاعل مع بعضها بعضاً.

تجربة ميدانية **Field experiment**

تجربة يجري تنفيذها في بيئة حدوثها حسب الشكل المتوقع

مذكرات ميدانية **Field notes**

السجلات المدونة للملاحظات والأحاديث أثناء مشروع بحث نوعي

اختبار ميداني **Field testing**

عملية اختبار أداة بحثية لمجموعة صغيرة من الأشخاص لتحديد عوامل الغموض/ الالتباس أو المشكلات الأخرى قبل إعداد الصيغة النهائية.

مقابلة مركّزة **Focused interview**

أسلوب في البحث النوعي لجمع البيانات يستخدم أسئلة مفتوحة ومرنة حيث يتمتع المستجيبون بحرية الإجابة والتعبير بكلماتهم الخاصة.

دراسة تتبعية **Follow-up study**

استقصاء التطور اللاحق للأفراد عقب معالجة أو برنامج.

توزيع تكراري **Frequency distribution**

تمثيل جدولي للدرجات التي حصلت عليها مجموعة من الأفراد.

Frequency polygon

راجع polygon.

اختبار فائي (F) **F-test**

إجراء إحصائي يستخدم لاختبار فرضيات حول الفروق بين وسطين أو أكثر، ولأغراض أخرى.

- G -

خطأ السماحة/ الكرم **Generosity error**

ميل مقدّر لمنح كل فرد فائدة الشك بإعطائه تقديرات عالية عندما لا يكون متيقناً

- H -

أثر/ ظاهرة الهالة **Halo effect**

ميل الانطباع العام لمقدّر، عن فرد، بالتأثير على تقديره الممنوح لجوانب متعددة أو أنواع من السلوك لذلك الفرد.

أثر/ ظاهرة هوثورن **Hawthorne effect**

تأثير على المتغير التابع ينشأ عن معرفة أفراد مجموعة المعالجة بأنهم يشاركون في تجربة

مدرّج تكراري **Histogram**

بيان/ مبيان يمثل التوزيع التكرار بأعمدة.

بحث تاريخي **Historical research**

محاولة منظمة لتأكيد الحقائق والوصول إلى نتائج حول الماضي

تاريخ **History**

أحداث أو ظروف، غير المتغير المستقل، تنتج تغيرات في المتغير التابع، وهي مشكلة صدق داخلي.

مجموعة متجانسة **Homogeneous group**

مجموعة من الأفراد الذين يكونون الشيء ذاته أو متماثلين حسب متغير ذي صلة.

اختبار متجانس **Homogeneous selection**

استخدام أفراد متماثلين بقدر الإمكان حسب خصائص ذات صلة بغية تقليص تأثير متغيرات دخيلة.

أداة بشرية **Human instrument**

الباحث كوسيلة لجمع المعلومات في بحث نوعي.

فرضية **Hypothesis**

مقترح أولي مؤقت يفترض كحل لمشكلة، فهي صياغة/ عبارة لتوقعات الباحث حول علاقة بين متغيرات الدراسة

اختبار الفرضية **Hypothesis test**

مجموعة من الملاحظات لتحديد ما إذا كانت هذه الملاحظات تؤيد أو تخفق في تأييد علاقة مفترضة.

- I -

استقراء غير تام **Imperfect induction**

عملية استدلال من عينة لمجموعة، إلى ما هو خاصة للمجموعة كلها.

تهديد/ مخاطر التنفيذ **Implementation threat**

تهديد للصدق الداخلي في دراسة، ينشأ عن التغيرات في تنفيذ معالجة.

استقراء **Induction**

اكتساب معرفة، انطلاقاً من دراسية منهجية/ منظمة وملاحظات حقائق خاصة أو سلسلة أحداث جزئية، والوصول بعدئذ إلى نتائج عامة.

تحليل استقرائي **Inductive analysis**

عملية البدء بملاحظات (بيانات) والمتابعة نحو صياغة فرضية وتنقيحها والوصول نهائيا إلى نظرية.

فرضية استقرائية **Inductive hypothesis**

فرضية تشتق خلال تعميم من ملاحظة.

إحصاء استدلالي **Inferential statistics**

أساليب تتيح للمرء صياغة تعميمات مؤقتة، من بيانات عينة، إلى المجتمع الأصلي المستمدة منه العينة.

Informed consent (right to) **(الحق في) الموافقة المطّلعة**

حق الفرد، في دراسة بحثية، في معرفة طبيعية وهدف الدراسة ولأن يقبل أو يرفض المشاركة.

Institutional review board (IRB) **مجلس رسمي/ مؤسساتي للمراجعة**

لجنة تقرر ما إذا كان مقترح البحث يلبي المعايير الفدرالية والقانونية والأخلاقية الأخرى.

Instrument **أداة**

وسيلة لتعريف متغير إجرائيا.

Instrumentation threat **تهديد استخدام الأدوات**

احتمال أن تعود نتائج دراسة إلى التغيرات في الطريقة التي تستخدم فيها الأدوات لتعريف المتغيرات إجرائياً، وهو تهديد للصدق الداخلي.

Interaction **تفاعل (تبادل التأثير)**

ناتج تجربة عاملية (factorial experiment) عندما تتغير تأثيرات متغير مستقل، على متغير تابع، حسب مستويات مختلفة لمتغير مستقل آخر.

Intercept **مقطع (تقاطع)**

القيمة التي يقطع فيها خط الارتداد محوراً. فمقطع Y هو النقطة في خط ارتداد Y على X عندما يقطع الخط المحور Y، أما مقطع X فهو النقطة في ارتداد X على Y حيث يقطع الخط المحور X.

Internal consistency **ثبات/ اتّساق داخلي**

إجراء لتقييم الثبات بقياس المدى الذي يكون فيه الارتباط المتبادل لبنود اختبار فرعي أو مقياس موجباً، وهكذا فإنها كلها تقيس المفهوم البنائي ذاته أو السمة ذاتها، فهو المدى الذي يقيس به اختبار ما خاصة واحدة فقط.

Internal criticism **نقد داخلي**

تقييم صدق وثيقة في بحث تاريخي

Internal validity **صدق داخلي**

المدى الذي تكون فيه الفروقات الملاحظة لمتغير تابع في تجربة ما، ناجمة عن المتغير المستقل، وليست عن متغير دخيل أو متغيرات لا تخضع للضبط.

Internet **شبكة الاتصال الحاسوبي (إنترنت)**

سلسلة شبكات الاتصال الحاسوبي المستخدمة لنشر المعلومات أو البحث عنها.

ثبات المقدّرين المتبادل **Interrater reliability**

المدى الذي يقدم فيه ملاحظان أو أكثر نتائج متماثلة عندما يلاحظان الفرد نفسه خلال الفترة الزمنية ذاتها.

سُلم/ مقياس فتري **Interval scale**

سلم قياس يرتب الأشياء أو الأحداث ونقاطه متساوية الأبعاد عن بعضها بعض.

مقابلة **Interview**

طرح أسئلة شفوية على فرد ما.

تحيز المقابل **Interviewer bias**

تحيز يطرأ في المقابلات عندما يعزز المقابل إيجابياً بصورة شفوية أو غير شفوية الاستجابات "الصحيحة" أو يعزز سلبيا الاستجابات الخاطئة.

جدول/ برنامج مقابلة **Interview schedule**

وثيقة يستخدمها المقابل، تحتوي على التعليمات والأسئلة في ترتيب محدد بالإضافة إلى عبارات انتقالية.

استبيان **Inventory**

مجموعة من العبارات يستجيب لها الأفراد بتوضيح ما إذا كانت العبارة تصفهم (تنطبق عليهم) أم لا، وتستخدم في تقييم الشخصية.

تحليل البنود/ المفردات **Item analysis**

تحليل اختبار يحدد عدد ونسبة الاستجابات الصحيحة لكل بند وارتباط درجات ذلك البند مع درجات الاختبار الكلية.

- J -

تأثير/ ظاهرة جون هنري **John Henry effect**

تأثير يطرأ عندما يكون أداء المجموعة الضابطة فوق متوسطها الاعتيادي عندما تدرك أنها في منافسة مع مجموعة تجريبية تستخدم طريقة أو إجراءاً جديدين.

- K -

معادلة كودر – ريتشاردسون **Kuder-Richardson formula**

معادلة لتقرير ثبات الاتساق الداخلي (التجانس) لأداة من تطبيق واحد لها دون تجزئة الاختبار إلى نصفين.

- L -

Laboratory experiment **تجربة معملية**

تجربة يجري تنفيذها في بيئة ذات ضبط عالٍ

Low of the single independent variable **قانون المتغير المستقل الواحد**

قاعدة تقول، إن التحديد غير الملتبس، لتأثير متغير مستقل على متغير تابع، يفرض أن يكون المتغير المستقل هو الفرق الوحيد بين المجموعتين التجريبية والضابطة قبل قياس المتغير التابع.

Level of significance **مستوى الدلالة**

الاحتمال الأكبر للخطأ المقبول في رفض الفرضية الصفرية، وهو عادة p=0.05 أو p=0.01 في البحوث التربوية.

Likert scale **مقياس ليكرت**

سلم قياس / مقياس يتكون من سلسلة من العبارات يتبعها خمسة أصناف استجابات، تمتد عادة من "أوافق بشدة/ جداً" إلى "لا أوافق بشدة/ جداً".

Likert-type item **بند حسب نمط ليكرت**

عبارة تماثل تلك التي في مقياس ليكرت حيث خيارات الاستجابات تكون على "متصل كمّي: Continuum" من "أوافق بشدة / جداً" إلى "لا أوافق بشدة / جداً".

Linear relationship **علاقة خطية**

علاقة ارتباط حيث تكون النقاط المحددة، في بيان التشتت/ الانتشار، لأزواج الدرجات تتواءم مع خط مستقيم.

Longitudinal study **مسح طولي**

مسح تجمع فيه البيانات عدة مرات طوال فترة ممتدة من الزمن.

- M -

Main effect **التأثير/ الأثر الرئيسي**

التأثير الذي يكون للمتغير المستقل على المتغير التابع في تجربة عاملية " Factorial experiment" بصرف النظر عن تأثير المتغيرات المستقلة الأخرى.

Manipulated variable **متغير مفعّل**

المتغير الذي يغيره/ يفعله القائم بالتجربة بصورة منهجية/ منتظمة لكي يلاحظ التأثير على متغير تابع.

هامش الخطأ **Marginal error**

تقدير المدى الذي يحتمل أن تنحرف فيه نتائج العينة عن قيمة المجتمع.

تناظر/ مزاوجة **Matching**

تحديد أزواج الأفراد الذين يتماثلون، قدر الإمكان، قبل بدء التجربة.

تهديد/ مخاطر النضج **Maturation threat**

إمكانية أن تعود النتائج الملاحظة في دراسة إلى تغيرات تطرأ على الأفراد كنتيجة لمرور الزمن، وليست كنتيجة للمتغير المستقل، وهي مشكلة صدق داخلي.

وسط/ متوسط حسابي **Mean**

مقياس للنزعة المركزية لتوزيع بيانات فترية، وهو خارج قسمة مجموع الدرجات على عدد الدرجات في التوزيع.

وسيط **median**

النقطة في توزيع التي تقع تحتها 50 في المائة من الدرجات (المئيني الخمسون) ويستخدم في بيانات الترتيب أو الفترة.

تعريف إجرائي مقيس **Measured operational definition**

تعريف يصاغ بدلالة الإجراءات التي يستخدمها الباحثون لقياس مفهوم.

كتب القياس العقلي **Mental Measurement Yearbooks**

سلسلة منشورات بقوائم الاختبارات المتوفرة، مع توصيفات ومطالعات نقدية لهذه الاختبارات.

ما وراء التحليل (تحليل التحليلات) **Meta-analysis**

الجمع المنهجي/ المنظم لبيانات كمية من عدد من الدراسات التي تتقصى العلاقة ذاتها بين المتغيرات.

منوال **Mode**

الدرجة التي تكون الأكثر تكراراً في توزيع للدرجات، ويستخدم مع البيانات الاسمية أو الترتيبية أو الفترية.

التحليل متعدد العوامل للتباين **Multifactor analysis of variance**

تحليل للتباين مع أكثر من متغير مستقل.

تصميم خط القاعدة المتعدد **Multiple-baseline design**

تجربة يتم فيها إجراء قياسات عدة متغيرات تابعة لخط قاعدة، لفرد واحد، وبعدئذ تبدأ

معالجات جرى تصميمها لإحداث تغييرات في المتغيرات التابعة خلال أوقات مختلفة لكل متغير تابع. وتستخدم أيضاً لتحديد تأثيرات/ نتائج المعالجة ذاتها، بالبدء فيها في أوقات مختلفة لأفراد آخرين.

ارتباط متعدد **Multiple correlation**

الارتباط بين مجموعة عوامل تنبؤ (predictors) ومتغير تابع واحد. ويشير إلى مدى العلاقة معامل الارتباط المتعدد (R).

ارتداد متعدد **Multiple regression**

التنبؤ لمعيار (متغير تابع: Criterion) باستخدام اثنين أو أكثر من متغيرات (عوامل) التنبؤ بضمّها معاً. ويوزن (يقدر) كل عامل تنبؤ بنسبة ترجيح لإسهامه في دقه التنبؤ. وتدعى المعادلة التي توضح الأوزان المخصّصة لكل عامل تنبؤ "معادلة الارتداد المتعدد".

- N -

استقصاء طبيعي **Naturalistic inquiry**

دراسة للأفراد في بيئتهم الاعتيادية دون فرضيات مقررة مسبقاً

محيط/ وضع طبيعي **Natural setting**

بيئة للبحث تقع فيها النشاطات حسب المساق الاعتيادي للحوادث، خلافاً لبيئة تكون فيها الحوادث مصطنعة أو مفعّلة من أجل الغرض للدراسة.

ارتباط سالب/ سلبي **Negative correlation**

ارتباط تتلازم فيه درجات عالية لمتغير مع درجات متدنية للمتغير الآخر.

منحن سالب الالتواء **Negatively skewed curve**

مضطلع تحتشد فيه العديد من الدرجات في الطرف الأعلى (اليمين) والدرجات المتدنية تنتشر نحو الطرف الأيسر (المتدني)، ويمتد ذيل المنحنى إلى اليسار.

سُلّم/ مقياس اسمي (نوعي) **Nominal scale**

سلم قياس يصنف الأشياء أو الأفراد إلى أصناف/ فئات تختلف نوعياً وليس كمياً.

فرضية لا اتجاهية **Nondirectional hypothesis**

فرضية تنص على وجود علاقة ستجري ملاحظتها/ رصدها بين متغيرات، لكنها لا تحدد الاتجاه المتوقع للاستنتاجات.

اختبار لا اتجاهي (اختبار ذو طرفين/ ذيلين) Nondirectional test (two-tailed-test)
اختبار إحصائي يأخذ بالاعتبار الفروق في كلا الاتجاهين (أكبر أو أصغر) من قيمة محددة بالفرضية الصفرية، فيتم رفض الفرضية الصفرية إذا كان الفرق كبيراً بما يكفي في "كلا" الذيلين لتوزيع معاينة للإحصاءة.

ملاحظ/ راصد غير مشارك Nonparticipant observer
طريقة بحث لا يشارك الملاحظ فيها بالوضع/ بالموقف قيد الدراسة، فيكون ملاحظاً فقط.

معاينة لا احتمالية Nonprobability sampling
معاينة لاختيار غير عشوائي.

عدم الاستجابة Nonresponse
موقف يتلقى فيه الأشخاص مسحاً، غير أنهم يخفقون في إعادة الأداة المستكملة.

منحنى اعتيادي/ اعتدالي/ نموذجي Normal curve
توزيع، افتراضي متناظر له شكل الجرس، للدرجات يستخدم كنموذج لتوزيعات عدة تقع بصورة طبيعية وفي عدة اختبارات إحصائية.

اختبار مرجعي المتوسط Norm-referenced test
اختبار يمكن للمرء من خلاله مقارنة أداء فرد مع أداء الآخرين الذين أجروا الاختبار.
(ملحوظة: تشير Norm إلى معاني مثل معيار، متوسط..." وتشير هنا إلى المتوسط (الوسط أو الوسيط" وقد أثرنا ترجمتها إلى "متوسط"حسبما يذكر واضعو هذا المصطلح".

فرضية صفرية Null hypothesis
فرضية تنص على عدم وجود تأثير، أو فرق، أو علاقة، بين متغيرات، فهي النقيض لفرضية البحث- وهي تطلّع/ أمل الباحث في رفضها.

- O -

اختبار ذو طرف/ ذيل واحد One-tailed test
أنظر Directional test

سؤال مفتوح (النهاية) Open-ended question
اختبار لا يحتوي على بدائل/ خيارات استجابات محددة، لكنه يتيح للمستجيب أن يردّ بالطريقة التي يختارها.

Operational definition **تعريف إجرائي**

تعريف يحدد إجراءاً أو عملية يتوجب اتباعها في إنتاج/ بناء أو قياس مفهوم.

Ordinal scale **سُلّم/ مقياس ترتيبي**

سلم قياس يرتب تدريجياً الأشياء أو الحوادث وفقاً لمدى ما تمتلكه من الخاصة قيد الاهتمام، غير أنه لا يسعنا الافتراض أن المسافات بين النقاط متساوية.

- P -

Panel study **دراسة المجموعة ذاتها**

دراسة مسح طولي حيث يتم جمع البيانات من العينة ذاتها للأفراد حسب نقاط زمنية مختلفة.

Parallel form **الشكل المماثل / المكافئ**

انظر Equivalent - form.

Parameter **معْلم / معْلمة، بارامتر**

خاصة لمجتمع إحصائي، مثل وسط المجتمع الإحصائي (μ) أو الانحراف المعياري للمجتمع الإحصائي (σ).

Parsimony principle **مبدأ الاقتصاد**

مبدأ ينصّ أنه إذا تساوت / تماثلت الأشياء الأخرى فإن التفسير الأبسط للظاهرة يُفضّل على التفسيرات الأكثر تعقيداً.

Partial control **ضبط / تحكّم جزئي**

بعض من الضبط وليس كله لمتغيرات دخيلة ذات صلة.

Partial correlation **ارتباط جزئي**

إجراء إحصائي لوصف العلاقة بين متغيرين في الارتباط بينهما مع متغير ثالث جرى استبعاده / تنحيته.

Participant observation **ملاحظة / رصد مشارك**

طريقة بحث يصبح فيها الباحث طرفا يشارك في نشاطات المجموعة أو الموقف قيد الدراسة.

Path analysis **تحليل المسار**

إجراء إحصائي لاستقصاء العلاقات العلّية بين متغيرات مرتبطة (correlated).

Pearson product moment coefficient (Pearson r)

معامل بيرسون التتابعي / لضرب العزوم (بيرسون r)

مؤشر ارتباط لبيانات فترية أو نسبية، وهو وسط حواصل ضرب الدرجات Z للمتغيرين.

Performance assessment **تقييم الأداء**

نوع من التقييم يتطلب من الأفراد أن يبينوا ما يعرفون بأدائهم واجبات محددة وليس حسب أداة ورقة وقلم.

Periodical index **فهرس الدوريات**

نشرة تصدر في أجزاء متتابعة لتوفير دليل منهجي لتحديد المعلومات في مجلات أو مصادر أخرى.

Phi coefficient **المعامل "فاي"**

معامل ارتباط يستخدم عندما يقاس كلاهما حسب سلّم ، مقياس اسمي.

Pilot study **دراسة استطلاعية / استكشافية**

محاولة تجريبية مع عدد قليل من الأفراد لتقييم الملاءمة والقابلية العملية للإجراءات وجمع بيانات الأدوات.

Placebo **عقار وهمي**

مادة خاملة كيميائياً أو محايدة تعطى للأفراد في تجربة لجعلهم يعتقدون أنهم يتلقون العقار أو المعالجة التجريبية.

Polygon **مضلع**

شكل يبين الدرجات وتكرارها بربط نقاط التقاطع لكل من الدرجات مع التكرارات.

Population **مجتمع إحصائي**

المجموعة الأكبر التي يهدف الباحث للتعميم عليها، وتشمل "كل" عناصر صنف من الناس أو الحوادث أو الأشياء.

Population validity **صدق تمثيل المجتمع الإحصائي**

درجة جودة تمثيل العينة لمجتمع إحصائي، فهو المدى الذي يمكن فيه تعميم نتائج دراسة على مجتمع إحصائي قيد الاهتمام.

Positive correlation **ارتباط موجب**

ارتباط تتلازم فيه الدرجات العليا لمتغير مع الدرجات العليا لمتغير آخر، والدرجات الدنيا تتلازم مع الدرجات الدنيا.

Positively skewed distribution **توزيع موجب الالتواء**

مضلع يبين احتشاد عديد من الدرجات قرب / في الطرف الأدنى (الأيسر) وتنتشر الدرجات العليا نحو الطرف الأعلى (الأيمن)، ويمتد الذيل نحو اليمين.

Power calculation **حساب / تقدير القوة**

إجراء رياضي لتحديد حجم العينة الضرورية لرفض الفرضية الصفرية حسب مستوى دلالة مقرر مع حجم تأثير مفروض.

Predictive validity **صدق تنبؤي**

راجع Criterion-related validity.

Predictor **عامل تنبؤ**

متغير يتم التنبؤ بواسطته في دراسة تنبؤية.

Preexeperimental research **البحث قبل التجريبي**

تصميمات تجريبية يكون فيها ضبط المتغيرات الدخيلة ضعيفاً أو معدوما، ومن ثم الصدق الداخلي ضعيف.

Pretest sensitization **حساسية الاختبار القبلي**

تأثير الاختبار القبلي على الأفراد الذي يدفعهم إلى الاستجابة على نحو مختلف للمعالجة، خلافاً لما قد يفعلوه بدون الاختبار القبلي.

Primary source **مصدر أولي / أصلي**

وثائق أو آثار أو بقايا أصلية أو السجلات الشهود عيان تستخدم في بحث تاريخي.

Probability sampling **معاينة احتمالية**

معاينة تستخدم الاختيار العشوائي، وهذا يعني أن لكل عنصر في المجموعة فرصة (احتمال) اختيار غير صفرية.

Probe **تحقيق / سبر دقيق**

سؤال متابعة يستخدم أثناء مقابلة لاستخلاص معلومات واضحة واستجابات أكثر دقة ممن تجرى مقابلته.

Projective technique **أسلوب إسقاطي**

طريقة لتقييم الشخصية بتحليل استجابات الفرد نحو مادة ملتبسة / غامضة.

Purposive sampling **معاينة قصدية / هادفة**

أسلوب معاينة غير احتمالية حيث يجرى الحكم على أفراد بكونهم ممثلين للمجتمع الإحصائي لاحتوائهم في العينة.

- Q -

Qualitative research **بحث نوعي**
مصطلح عام لمجموعة منوّعة من الأساليب البحثية التي تدرس الظواهر في بيئتها الطبيعية، دون فرضيات مقررة مسبقاً.

quantitative research **بحث كمي**
استقصاء يستخدم تعريفات إجرائية لتوليد بيانات عددية بهدف الإجابة على فرضيات أو أسئلة مقررة مسبقاً

Quartile deviation **انحراف رُبيْعي / إرباعي**
مقياس ترتيبي للتشتت / الانتشار، منتصف المسافة بين الإرباعين الأول والثالث.

Quartiles **الرُّبيْعات ، الإرباعات**
نقاط الدرجات الثلاث التي تقسم توزيعاً تكرارياً إلى أربعة أرباع مع عدد متساو من الدرجات في كل منها. الوسيط هو الإرباع الثاني.

Quasi - experimental research **بحث شبه تجريبي**
بحث يكون بمقدور الباحث في ضبط المعالجة ومقياس المتغير التابع، غير أنه ليس بمقدوره ضبط تخصيص الأفراد للمعالجة.

Questionnaire **استبيان**
أداة يوفر المستجيبون فيها استجابات مكتوبة لأسئلة أو وضع إشارة على بنود لتوضيح استجاباتهم.

Quota sampling **معاينة حصصية**
أسلوب معاينة غير احتمالية تحدد حجم كل مجموعة جزئية ذات صلة بالمجتمع الإحصائي، ويجرى بعدئذ اختيار غير عشوائي للأفراد لتكوين عينة تكون فيها نسبة كل مجموعة جزئية هي نفسها في العينة كما هي في المجتمع الإحصائي.

- R -

Random assignment (randomization) **التخصيص العشوائي/ تعْشية)**
تخصيص أفراد عينة إلى مجموعة تجريبية أو ضابطة بطريقة عشوائية.

Randomized matching **تناظر / تزاوج عشوائي**
إجراء / أسلوب يتم فيه أولا مزاوجة الأفراد حسب متغيرات ذات صلة، ويجرى بعدئذ تخصيص عشوائي لفرد في كل زوج للمجموعة التجريبية، والآخر للمجموعة الضابطة.

Random sample **عينة عشوائية**

عينة يتم اختيارها بفعل الصدفة (عشوائياً) بحيث يكون لكل فرد في المجتمع، الاحتمال نفسه، في اختياره.

Random selection **اختيار عشوائي**

عملية اختيار عينة عشوائياً بحيث يكون احتمال اختيار كل فرد من المجتمع متساو للجميع.

Range **مدى**

مقياس اسمي للتشتت، الفرق بين أعلى وأدنى الدرجات مضافا له " 1 " من وحدات القياس.

Ranking item **بند ترتيبي**

نمط من سؤال مسح يطلب إلى المستجيبين تحديد ترتيب تفضيلهم/أرجحيتهم بين عدد الخيارات.

Ratio scale **سُلّم / مقياس نسبي**

سلّم قياس يوفر نقطة صفر حقيقي وفترات متساوية كذلك.

Regression (statistical) **ارتداد (إحصائي)**

النـــزعة الإحصائية للدرجات المتطرفة، في القياس الأول، إلى التحرك أقرب نحو الوسط في القياس الثاني.

Regression equation **معادلة الارتداد**

معادلة تشتق إحصائياً لتنبؤ قيمة متغير تابع من قيم متغيرات مستقلة.

Regression line **خط الارتداد**

الخط "الأفضل مطابقة / توفيقاً" لمجموعة من الدرجات محددة في بيان للتشتت.

Regression threat **تهديد / مخاطر الارتداد**

مشكلة الصدق الداخلي التي تنشأ عندما تعود نتائج دراسة إلى نزعة المجموعات، التي تم اختيارها على أساس الدرجات المتطرفة، لتتحرك (لترتد) نحو المتوسط في قياسات لاحقة، بغض النظر عن المعالجة التجريبية.

Reliability **ثبات**

المدى الذي يتيح المقياس فيه نتائج ثابتة / مستقرة، أي المدى الذي تتحرر فيه الدرجات من الخطأ العشوائي.

Research proposal **مقترح (خطة) بحث**

خطة متدرجة (خطوة فخطوة) لتنفيذ وإتمام دراسة بحثية.

اتجاه الاستجابة **Response set**

طريقة تحكمها العادة في الاستجابة، بمعنى أنها مستقلة عن المحتوى لمسح خاص أو بند اختبار، كأن يضع المرء دوماً إشارة أمام "غير مقرر" بصرف النظر عن العبارة.

العلّية العكسية **Reverse causality**

فرضية بديلة تنص أن المتغير التابع المفترض هو في الحقيقة المتغير المستقل (ص تتسبّب في س وليست س هي التي تتسبّب في ص).

RIE (مصادر في التربية) **RIE (Resources in Education)**

مجلة ERIC الشهرية التي تفهرس البحوث التي ترعاها حكومة الولايات المتحدة بالإضافة إلى مواد متفرقة وتوفر ملخصات وجيزة عن هذه الوثائق.

- S -

عيّنة **Sample**

مجموعة يجرى اختيارها من مجتمع إحصائي للملاحظة في دراسة ما.

مسح عيّنة **Sample survey**

مسح يجرى فقط على جزء (عينة) من المجتمع، يستخدم عادة استبياناً أو مقابلة.

سلّم / مقياس **Scale**

تدريج (مقياس مدرّج) تكون له وحدات كمية، في العادة، يقيس الدرجة (المدى) التي يبديها الأفراد لنـــزعات أو خصائص محددة.

بيان / مبيان التشتت **Scattergram**

شكل يستخدم نقاطا بيانية محددة تمثل تقاطع درجتي الفرد (X, Y) وتخدم لتوضيح علاقة بين المتغيرين X، Y.

اختبار الاستعداد الدراسي **Scholastic aptitude test**

اختبار يقيس القدرات التي تتنبأ في النجاح بالمطالب الأكاديمية.

المنهج العلمي **Scientific approach**

طريقة لبحث عن المعرفة تتضمن كلا من التفكيرين الاستقرائي والاستنتاجي لبناء فرضيات تكون عرضة لاختبار موضوعي شديد الدقة.

مصدر ثانوي **Secondary source**

معلومات غير أولية (Secondhand) مثل وصف لحدث تاريخي كتبه شخص آخر غير شاهد عيان.

تفاعل الاختيار والنضوج **Selection - maturation interaction**

هذه مشكلة صدق داخلي تنشأ عندما يجرى اختيار الأفراد بطريقة تنضج فيها المجموعتان التجريبية والضابطة حسب معدلات مختلفة.

سلم / مقياس تمايز المعاني **Semantic differential**

طريقة تقييم لاتجاهات الأفراد بالطلب إليهم وضع إشارة لنقاط على تدريج بين صفات ثنائية القطب.

بحث الفرد الواحد **Single-Subject research**

تجربة ذات فرد واحد أو مجموعة تامة.

ميل (مستقيم) **slope**

في خط انحدار، يحدد الميل مقدار متوسط التغير في متغير Y لكل وحدة تغير في X.

تحيز الرغبة الاجتماعية **Social desirability bias**

تحيز يطرأ في المقابلات عندما يعطى الأفراد استجابات لتعزيز صورهم بدلا من استجابات أمينة / غير متحيزة.

مبيان اجتماعي **Sociogram**

شكل يستخدم الأسهم لتوضيح أنماط العلاقات الاجتماعية بين الأفراد في مجموعة.

أسلوب القياس الاجتماعي **Sociometric technique**

دراسة بنية اجتماعية بالطلب إلى كل فرد في المجموعة اختيار أفراد آخرين حسب معيار محدد.

تصميم سلومون ذو المجموعات الأربع **Solomon four-group design**

تصميم تجريبي ينطوي على تخصيص عشوائي لأفراد كل من أربع مجموعات حيث تخضع مجموعتان لاختبار قبلي ولا تخضع المجموعتان الباقيتان له، ثم تتلقى المعالجة مجموعة، خضعت للقبلي ومجموعة أخرى لم تخضع له، وأخيرا تخضع المجموعات الأربع لاختبار بعدي.

تصميم سلومون ذو المجموعات الثلاث **Solomon three-group design**

تصميم تجريبي ينطوي على تخصيص عشوائي لأفراد كل من ثلاث مجموعات: (1) اختبار قبلي + معالجة، (2) اختبار قبلي + لامعالجة، أو (3) لا اختبار قبلي + معالجة.

معادلة سبيرمان -براون **Spearman - Brown formula**

إجراء إحصائي يستخدم ارتباط التجزئة النصفية لاختبار لتقدير ثبات الاختبار الكلي.

"رو" سبيرمان **Spearman rho**

مؤشر ارتباط يستخدم مع البيانات الترتيبية.

ثبات التجزئة النصفية **Split-half reliability**

نوع من ثبات الاستقرار الداخلي يتم الحصول عليه بتجزئة مصطنعة لاختبار إلى نصفين وإقامة ارتباط بين درجات الأفراد في نصفي الاختبار.

نتيجة زائفة / صورية **Spurious result**

علاقة ظاهرية بين متغيرين لا تكون حقيقية.

انحراف معياري **Standard deviation**

مقياس مدى انحراف درجات فرد عن وسط توزيع، وهو الجذر التربيعي للتباين الذي يعتبر مقياساً للتشتت ويستخدم مع البيانات الفترية.

الخطأ المعياري للتقدير **Standard error of estimate**

مقياس للفرق المتوقع بين الدرجات المتنبأ بها والدرجات الحقيقية.

الخطأ المعياري للقياس **Standard error of measurement**

مؤشر لمقدار خطأ القياس في درجات اختبار، ونظريا، هو الانحراف المعياري لتوزيع الدرجات الملاحظة / المرصودة حول الدرجة الحقيقية للفرد.

الخطأ المعياري للوسط **Standard error of the mean**

الانحراف المعياري لأخطاء معاينات الوسط، ويشير إلى مقدار ما يسعنا توقعه لاختلاف أوساط العينات العشوائية، المستمدة من مجتمع إحصائي واحد، عن طريق الصدفة وحدها.

اختبار مقنّن **Standardized test**

اختبار ذو محتوى مميّز، يحدد طريقة إجرائه وتصحيحه، وتشتق معلومات متوسطاته (معاييره) وثباته وصدقه بتطبيقه على عينات ممثلة.

درجة معيارية **Standard Score**

درجة محوّلة، يجرى التعبير عنها بعدد الوحدات المعيارية لبعد هذه الدرجة عن الوسط، ويحدد معيار اختياري الوسط والانحراف المعياري للدرجات المعيارية. فالدرجة (المعيارية) Z، مثلا، لها وسط (صفر) وانحراف معياري (1)، وتستخدم الدرجات T وسطا (50) وانحرافا معياريا (10).

مقارنة المجموعة الثابتة **Static group comparison**

تصميم قبل تجريبي، ينطوي على مجموعتين غير متكافئتين على الأقل، حيث تتلقى المعالجة مجموعة واحدة وتجرى كلاهما اختباراً بعديا. ويوفر ضبط قليلاً أو معدوما.

Statistic **إحصاءة**

خاصة للعينة.

Statistical equivalence **تكافؤ إحصائي**

حالة تقع عندما يعود أي فرق بين مجموعات، للصدفة وحدها، وينشأ عن التخصيص العشوائي للأفراد في المجموعات، مجموعات متكافئة إحصائياً.

Statistically significant result **نتيجة دالة إحصائيا**

نتيجة يكون احتمال حدوثها بفعل الصدفة أقل من احتمال محدد.

Stratified sampling **معاينة طبقية**

أسلوب معاينة احتمالية يجزئ المجتمع أولاً إلى مجموعات جزئية حسب متغيرات ذات صلة، مثل العمر، أو الحالة الاجتماعية، أو التعليم، ويقوم بعدئذ باختيار عشوائي للأفراد من كل مجموعة جزئية.

Style manual **كتيب / كراسة النمط**

كتيب يصف بتفصيل، شكل ونمط الرسائل والأطروحات والأعمال الرسمية المكتوبة الأخرى.

Subject **فرد**

عنصر في دراسة ما.

Summated rating scale **مقياس التقدير الجمعي**

سلم / مقياس يتوصل إلى الدرجة الكلية للشخص بإيجاد مجموع الاستجابات الموزونة / المرجحة لكل بنود المقياس – انظر Likert scale.

Survey **مسح**

دراسة عينة لاستقصاء مدى توزيع المتغيرات.

Symmetrical Sampling **معاينة متناظرة**

توزيع يمكن تمثيله بمضلع يكون جانبه الأيمن صورة مناظرة (عاكسة) لجانبه الأيسر.

Systematic sampling **معاينة منتظمة**

معاينة احتمالية حيث يتم اختيار العنصر ذي الرتبة k، للعينة، في قائمة المجتمع الإحصائي.

- T -

مجتمع مستهدف **Target population**

المجموعة الكلية (المجتمع) الذي يود الباحث تعميم نتائج الدراسة عليه.

اختبار **Test**

إجراء منظم لقياس عينة من السلوك.

اختبار الدلالة **Test of significance**

اختبار إحصائي يُستخدم لتقرير ما إذا كان يُحتمل أن تكون النتائج المستحصلة من فعل الصدفة. ويُستخدم لتقييم مصداقية الفرضية الصفرية.

ثبات إعادة الاختبار **test-retest reliability**

مقياس للثبات يتم الحصول عليه بإجراء ارتباط الدرجات من تطبيقين للمقياس ذاته على الأفراد ذاتهم.

نظرية **Theory**

مجموعة من القضايا أو الفرضيات المتداخلة معا، تمثل تفسيراً لظاهرة ما.

مقياس ثيرستون **Thurstone scale**

مقياس اتجاهات يحتوي على سلسلة من العبارات، لكل منها قيمة محددة مسبقا في المقياس، ويستجيب الأفراد بالإشارة إلى العبارات الأقرب مطابقة لاتجاهاتهم.

تصميم السلسلة الزمنية **Time-series design**

تجربة يجرى فيها قياس المتغير التابع في أوقات متعددة قبل وبعد المعالجة.

دراسة النـزعة / الاتجاه / التيار **Trend study**

مسح طولي تُجمع فيه البيانات من عينات مختلفة للمجتمع الإحصائي العام، في أوقات مختلفة لاستقصاء التغيرات مع الزمن.

مسح مثلثي **Triangulation**

تأكيد البيانات باستخدام إجراءات جمع بيانات متعددة، أو باستخدام مصادر متعددة للبيانات، ويستخدم في البحث النوعي.

تصميم تجريبي حقيقي **True experimental design**

تصميم تجريبي يقوم فيه الباحث بضبط المعالجة كما يستخدم العشوائية في تخصيص الأفراد للمعالجات.

الدرجة T **T-score**

درجة معيارية مع وسط (50) وانحراف معياري (10).

اختبار "ت – t" **t-test**

إجراء إحصائي لاختبار فرضية تتعلق بالفرق بين وسطين، ويستخدم، أيضاً، لأغراض أخرى.

اختبار في طرفين / ذيلين **Two-tailed test**

راجع Nondirectional test.

خطأ النوع الأول (I) **Type I error**

الخطأ الذي يقع عندما يرفض الباحث فرضية صفرية تكون، في الواقع، صحيحة.

خطأ النوع (II) **Type II error**

الخطأ الذي يقع عندما يخفق الباحث في رفض فرضية صفرية تكون، في الواقع، خاطئة.

- V -

صدق / صحة **Validity**

المدى الذي يستطيع فيه مقياس إبراز (استخلاص) المفهوم الأساسي الذي يزعم أنه يقيسه.

حالة / نـــزعة التغير – التشار / تشتت **variability**

التشتت أو الانتشار في توزيع للدرجات.

متغير **Variable**

تمثيل لمفهوم أو مفهوم بنائي، يمكنه أن يأخذ قيما لمجال. مثلاً، الطول، ودرجات اختبار القراءة، والاستعداد والجنس هي متغيرات.

تباين **Variance**

وسط مربعات انحراف الدرجات، فهو مقياس فتري لانتشار / تشتت الدرجات حول الوسط.

- Z -

الدرجة (المعيارية) Z **Z-score**

درجة معيارية تشير إلى مقدار بُعد درجة عن درجة الوسط بدلالة وحدات الانحراف المعياري.

اختبار " Z " **Z-test**

اختبار إحصائي يستخدم عندما يتطلب الأمر مقارنة وسط العينة مع وسط المجتمع الإحصائي ويكون الانحراف المعياري للمجتمع الإحصائي معلوماً.


دار
المسيرة
للنشر والتوزيع والطباعة
شركة جمال أحمد محمد حيف وإخوانه
www.massira.jo



شركة جمال أحمد محمد حيف وإخوانه
www.massira.jo

INTRODUCTION TO RESEARCH IN EDUCATION

مقدمة للبحث في التربية

